

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
1 August 2002 (01.08.2002)

PCT

(10) International Publication Number  
**WO 02/059321 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/40,  
15/51, 5/10, C07K 14/18, C12N 15/85

Pomezia (Rome) (IT). MIGLIACCIO, Giovanni [IT/IT];  
IRBM, Via Pontina Km. 30, 600, I-00040 Pomezia (Rome)  
(IT). PAONESSA, Giacomo [IT/IT]; IRBM, Via Pontina  
Km. 30, 600, I-00040 Pomezia (Rome) (IT).

(21) International Application Number: PCT/EP02/00526

(22) International Filing Date: 16 January 2002 (16.01.2002)

(74) Agent: THOMPSON, John; Merck & Co., Inc., European  
Patent Department, Terlings Park, Eastwick Road, Harlow,  
Essex CM20 2QR (GB).

(25) Filing Language: English

(26) Publication Language: English

(81) Designated States (*national*): CA, JP, US.

(30) Priority Data:  
60/263,479 23 January 2001 (23.01.2001) US

(84) Designated States (*regional*): European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE, TR).

(71) Applicant (*for all designated States except US*): ISTI-  
TUTO DI RICERCHE DI BIOLOGIA MOLECO-  
LARE P. ANGELETTI SPA [IT/IT]; Via Pontina Km.  
30, 600, I-00040 Pomezia (Rome) (IT).

Published:

— *without international search report and to be republished  
upon receipt of that report*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): DE FRANCESCO,  
Raffaele [IT/IT]; IRBM, Via Pontina Km. 30, 600, I-00040

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*



WO 02/059321 A2

(54) Title: HEPATITIS C VIRUS REPLICONS AND REPLICON ENHANCED CELLS

(57) Abstract: The present invention features nucleic acid containing one or more adaptive mutations, and HCV replicon enhanced cells. Adaptive mutations are mutations that enhance HCV replicon activity. HCV replicon enhanced cells are cells having an increased ability to maintain an HCV replicon.

## TITLE OF THE INVENTION

## HEPATITIS C VIRUS REPLICONS AND REPLICON ENHANCED CELLS

## CROSS-REFERENCE TO RELATED APPLICATIONS

- 5                   The present application claims priority to U.S. Serial No. 60/263,479, filed January 23, 2001, hereby incorporated by reference herein.

## BACKGROUND OF THE INVENTION

- 10                   The references cited in the present application are not admitted to be prior art to the claimed invention.

- It is estimated that about 3% of the world's population are infected with the Hepatitis C virus (HCV). (Wasley, *et al.*, 2000. *Semin. Liver Dis.* 20, 1-16.) Exposure to HCV results in an overt acute disease in a small percentage of cases, while in most instances the virus establishes a chronic infection causing liver inflammation and slowly progresses into liver failure and cirrhosis. (Iwarson, 1994. *FEMS Microbiol. Rev.* 14, 201-204.) In addition, epidemiological surveys indicate an important role of HCV in the pathogenesis of hepatocellular carcinoma. (Kew, 1994. *FEMS Microbiol. Rev.* 14, 211-220, Alter, 1995. *Blood* 85, 1681-1695.)

- 20                   The HCV genome consists of a single strand RNA of about 9.5 kb in length, encoding a precursor polyprotein of about 3000 amino acids. (Choo, *et al.*, 1989. *Science* 244, 362-364, Choo, *et al.*, 1989. *Science* 244, 359-362, Takamizawa, *et al.*, 1991. *J. Virol.* 65, 1105-1113.) The HCV polyprotein contains the viral proteins in the order: C-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B.

- 25                   Individual viral proteins are produced by proteolysis of the HCV polyprotein. Host cell proteases release the putative structural proteins C, E1, E2, and p7, and create the N-terminus of NS2 at amino acid 810. (Mizushima, *et al.*, 1994. *J. Virol.* 68, 2731-2734, Hijikata, *et al.*, 1993. *P.N.A.S. USA* 90, 10773-10777.)

- The non-structural proteins NS3, NS4A, NS4B, NS5A and NS5B presumably form the virus replication machinery and are released from the polyprotein. A zinc-dependent protease associated with NS2 and the N-terminus of NS3 is responsible for cleavage between NS2 and NS3. (Grakoui, *et al.*, 1993. *J. Virol.* 67, 1385-1395, Hijikata, *et al.*, 1993. *P.N.A.S. USA* 90, 10773-10777.) A distinct serine protease located in the N-terminal domain of NS3 is responsible for proteolytic cleavages at the NS3/NS4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B junctions. (Barthenschlager, *et al.*, 1993. *J. Virol.* 67, 3835-3844, Grakoui, *et al.*,
- 35

1993. *Proc. Natl. Acad. Sci. USA* 90, 10583-10587, Tomei, *et al.*, 1993. *J. Virol.* 67, 4017-4026.) NS4A provides a cofactor for NS3 activity. (Failla, *et al.*, *J. Virol.* 1994. 68, 3753-3760, De Francesco, *et al.*, U.S. Patent No. 5,739,002.) NS5A is a highly phosphorylated protein concurring interferon resistance. (De Francesco, *et al.*, 2000. *Semin Liver Dis.*, 20(1), 69-83, Pawlotsky, 1999. *J. Viral Hepat. Suppl.* 1, 47-48.)
- 5 NS5B provides an RNA polymerase. (De Francesco, *et al.*, International Publication Number WO 96/37619, Behrens, *et al.*, 1996. *EMBO* 15, 12-22, Lohmann, *et al.*, 1998. *Virology* 249, 108-118.)
- Lohmann, *et al.*, *Science* 285, 110-113, 1999, illustrates the ability of a
- 10 biscistronic HCV replicon to replicate in a hepatoma cell line. The biscistronic HCV replicon contained a neomycin cistron and an NS2-NS5B or an NS3-NS5B cistron. "NS2-NS5B" refers to a NS2-NS3-NS4A-NS4B-NS5A-NS5B polyprotein. "NS3-NS5B" refers to a NS3-NS4A-NS4B-NS5A-NS5B polyprotein.
- Bartenschlager, European Patent Application 1 043 399, published
- 15 October 11, 2000 (not admitted to be prior art to the claimed invention), describes a cell culture system for autonomous HCV RNA replication and protein expression. Replication and protein expression is indicated to occur in sufficiently large amounts for quantitative determination. European Patent Application 1 043 399 indicates that prior cell lines or primary cell cultures infected with HCV do not provide favorable
- 20 circumstances for detecting HCV replication.

## SUMMARY OF THE INVENTION

- The present invention features nucleic acid containing one or more adaptive mutations, and HCV replicon enhanced cells. Adaptive mutations are
- 25 mutations that enhance HCV replicon activity. HCV replicon enhanced cells are cells having an increased ability to maintain an HCV replicon.

- An HCV replicon is an RNA molecule able to autonomously replicate in a cultured cell and produce detectable levels of one or more HCV proteins. The basic subunit of an HCV replicon encodes for a HCV NS3-NS5B polyprotein along
- 30 with a suitable 5' UTR-partial core (PC) region and 3' UTR. The 5' UTR-PC region is made up of a 5'UTR region and about 36 nucleotides of the beginning of the core. Additional regions may be present including those coding for HCV proteins or elements such as the complete core, E1, E2, p7 or NS2; and those coding for other types of proteins or elements such as a encephalomyocarditis virus (EMCV) internal
- 35 ribosome entry site (IRES), a reporter protein or a selection protein.

The present application identifies different adaptive mutations that enhance HCV replicon activity. Enhancing replicon activity brings about at least one of the following: an increase in replicon maintenance in a cell, an increase in replicon replication, and an increase in replicon protein expression.

5            Adaptive mutations are described herein by identifying the location of the adaptive mutation with respect to a reference sequence present in a particular region. Based on the provided reference sequence, the same adaptive mutation can be produced in corresponding locations of equivalent regions having an amino acid sequence different than the reference sequence. Equivalent regions have the same  
10           function or encode for a polypeptide having the same function.

             Replicon enhanced cells are a preferred host for the insertion and expression of an HCV replicon. Replicon enhanced cells are initially produced by creating a cell containing a HCV replicon and then curing the cell of the replicon. The term "replicon enhanced cell" includes cells cured of HCV replicons and progeny  
15           of such cells.

             Thus, a first aspect of the present invention describes a nucleic acid molecule comprising at least one of the following regions: an altered NS3 encoding region, an altered NS5A encoding region, and an altered EMCV IRES region. The altered region contains one or more adaptive mutations. Reference to the presence of  
20           particular adaptive mutation(s) does not exclude other mutations or adaptive mutations from being present. Adaptive mutations are described with reference to either an encoded amino acid sequence or a nucleic acid sequence.

             A nucleic acid molecule can be single-stranded or part of a double strand, and can be RNA or DNA. Depending upon the structure of the nucleic acid  
25           molecule, the molecule may be used as a replicon or in the production of a replicon. For example, single-stranded RNA having the proper regions can be a replicon, while double-stranded DNA that includes the complement of a sequence coding for a replicon or replicon intermediate may useful in the production of the replicon or replicon intermediate.

30           Preferred nucleic acid molecules are those containing region(s) from SEQ. ID. NOs. 1, 2, or 3, or the RNA version thereof, with one or more adaptive mutations. Reference to "the RNA version thereof" indicates a ribose backbone and the presence of uracil instead of thymine.

             The presence of a region containing an adaptive mutation indicates that  
35           at least one such region is present. In different embodiments, for example, adaptive

mutations described herein are present at least in the NS3 region, in the NS5A region, in the NS3 and NS5A regions, in the EMCV IRES and NS3 regions, in the EMCV and NS5A regions, and in the EMCV IRES, NS3 and NS5A regions.

Another aspect of the present invention describes an expression vector  
5 comprising a nucleotide sequence of an HCV replicon or replicon intermediate coupled to an exogenous promoter. Reference to a nucleotide sequence "coupled to an exogenous promoter" indicates the presence and positioning of an RNA promoter such that it can mediate transcription of the nucleotide sequence and that the promoter is not naturally associated with the nucleotide sequence being transcribed. The  
10 expression vector can be used to produce RNA replicons.

Another aspect of the present invention describes a recombinant human hepatoma cell. Reference to a recombinant cell includes an initially produced cell and progeny thereof.

Another aspect of the present invention describes a method of making  
15 a HCV replicon enhanced cell. The method involves the steps of: (a) introducing and maintaining an HCV replicon into a cell and (b) curing the cell of the HCV replicon.

Another aspect of the present invention describes an HCV replicon enhanced cell made by a process comprising the steps of: (a) introducing and maintaining an HCV replicon into a cell and (b) curing the cell of the HCV replicon.

Another aspect of the present invention describes a method of making  
20 a HCV replicon enhanced cell comprising an HCV replicon. The method involves (a) introducing and maintaining a first HCV replicon into a cell, (b) curing the cell of the replicon, and (c) introducing and maintaining a second replicon into the cured cell, where the second replicon may be the same or different as the first replicon.

Another aspect of the present invention describes an HCV replicon enhanced cell containing a HCV replicon made by the process involving the step of introducing an HCV replicon into an HCV replicon enhanced cell. The HCV replicon introduced into the HCV replicon enhanced cell may be the same or different than the HCV replicon used to produce the HCV replicon enhanced cell. In a preferred  
25 embodiment, the HCV replicon introduced into an HCV replicon enhanced cell is the same replicon as was used to produce the enhanced cell.  
30

Another aspect of the present invention describes a method of measuring the ability of a compound to affect HCV activity using an HCV replicon comprising an adaptive mutation described herein. The method involves providing a  
35 compound to a cell comprising the HCV replicon and measuring the ability of the

compound to affect one or more replicon activities as a measure of the effect on HCV activity.

Another aspect of the present invention describes a method of measuring the ability of a compound to affect HCV activity using an HCV replicon enhanced cell that comprises an HCV replicon. The method involves providing a  
 5 compound to the cell and measuring the ability of the compound to effect one or more replicon activities as a measure of the effect on HCV activity.

Other features and advantages of the present invention are apparent from the additional descriptions provided herein including the different examples.  
 10 The provided examples illustrate different components and methodology useful in practicing the present invention. The examples do not limit the claimed invention. Based on the present disclosure the skilled artisan can identify and employ other components and methodology useful for practicing the present invention.

## 15 BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1G illustrate the nucleic acid sequence for the pHCVNeo.17 coding strand (SEQ. ID. NO. 3). The different regions of pHCVNeo.17 are provided as follows:

- 1-341: HCV 5' non-translated region, drives translation of the core-neo fusion protein;
- 20 342-1181: Core-neo fusion protein, selectable marker;
- 1190-1800: Internal ribosome entry site of the encephalomyocarditis virus, drives translation of the HCV NS region;
- 1801-7755: HCV polyprotein from non-structural protein 3 to non-structural protein 5B;
- 25 1801-3696: Non-structural protein 3 (NS3), HCV NS3 protease/helicase;
- 3697-3858: Non-structural protein 4A (NS4A), NS3 protease cofactor;
- 3859-4641: Non-structural protein 4B (NS4B);
- 4642-5982: Non-structural protein 5A (NS5A);
- 5983-7755: Non-structural protein 5B (NS5B); RNA-dependent RNA polymerase
- 30 7759-7989: HCV 3' non-translated region; and
- 7990-10690 plasmid sequences comprising origin of replication, beta lactamase coding sequence, and T7 promoter.

## DETAILED DESCRIPTION OF THE INVENTION

HCV replicons and HCV replicon enhanced cells can be used to produce a cell culture providing detectable levels of HCV RNA and HCV protein. HCV replicons and HCV replicon enhanced hosts can both be obtained by selecting  
5 for the ability to maintain an HCV replicon in a cell. As illustrated in the examples provided below, adaptive mutations present in HCV replicons and host cells can both assist replicon maintenance in a cell.

The detectable replication and expression of HCV RNA in a cell culture system has a variety of different uses including being used to study HCV  
10 replication and expression, to study HCV and host cell interactions, to produce HCV RNA, to produce HCV proteins, and to provide a system for measuring the ability of a compound to modulate one or more HCV activities.

Preferred cells for use with a HCV replicon are Huh-7 cells and Huh-7 derived cells. "Huh-7 derived cells" are cell produced starting with Huh-7 cells and  
15 introducing one or more phenotypic and/or genotypic modifications.

### Adaptive Mutations

Adaptive mutations enhance the ability of an HCV replicon to be maintained and expressed in a host cell. Adaptive mutations can be initially selected  
20 for using a wild type HCV RNA construct or a mutated HCV replicon. Initial selection involves providing HCV replicons to cells and identifying clones containing a replicon.

Nucleic acid sequences of identified HCV replicons can be determined using standard sequencing techniques. Comparing the sequence of input HCV  
25 constructs and selected constructs provides the location of mutations. The effect of particular mutation(s) can be measured by, for example, producing a construct to contain particular mutation(s) and measuring the effect of these mutation(s). Suitable control constructs for comparison purposes include wild type constructs and constructs previously evaluated.

Adaptive mutations were predominantly found in the HCV NS3 and NS5A regions. With the exception of two silent mutations in NS5A and NS5B, consensus mutations occurring in the NS region resulted in changes to the deduced amino acid sequence. Noticeably, the amino acid changes occurred in residues that are conserved in all or a large number of natural HCV isolates. HCV sequences are  
35 well known in the art and can be found, for example, in GenBank.

Adaptive mutations described herein can be identified with respect to a reference sequence. The reference sequence provides the location of the adaptive mutation in, for example, the NS3 or NS5A RNA, cDNA, or amino acid sequence. The remainder of the sequence encodes for a functional protein that may have the same, or a different, sequence than the reference sequence.

- Preferred NS3 and NS5A adaptive mutations and examples of changes that can be made to produce such mutations are shown in Tables 1 and 2. The amino acid numbering shown in Tables 1 and 2 is with respect to SEQ. ID. NO. 1. The nucleotide numbering shown in Tables 1 and 2 is with respect to SEQ. ID. NO. 2.
- SEQ. ID. NO. 1 provides the amino acid sequence of the Con1 HCV isolate (Accession Number AJ238799). SEQ. ID. NO. 2 provides the nucleic acid sequence of the Con1 HCV isolate.

TABLE 1

15

| Preferred NS3 Adaptive Mutations |            |
|----------------------------------|------------|
| Amino Acid                       | Nucleotide |
| gly1095ala                       | G3625C     |
| glu1202gly                       | A3946G     |
| ala1347thr                       | G4380A     |

TABLE 2

| Preferred NS5A Adaptive Mutations |               |
|-----------------------------------|---------------|
| Amino Acid                        | Nucleotide    |
| Lys@2039                          | AAA@6458      |
| asn2041thr                        | A6463C        |
| ser2173phe                        | C6859T        |
| ser2197phe                        | C6931T        |
| leu2198ser                        | T6934C        |
| ala2199thr                        | G6936A        |
| ser2204arg                        | C6953A (or G) |

"@" refers to an addition.



Preferred adaptive mutations identified with respect to a reference sequence can be produced changing the encoding region of SEQ. ID. NO. 1, or an equivalent sequence, to result in the indicated change. Preferred adaptive mutations provided in Tables 1 and 2 occur in amino acids conserved among different HCV isolates.

Adaptive mutations have different effects. Some mutations alone, or in combination with other mutations, enhance HCV replicon activity. In some cases, two or more mutations led to synergistic effects and in one case, a slightly antagonistic effect was observed.

An adaptive mutation once identified can be introduced into a starting construct using standard genetic techniques. Examples of such techniques are provided by Ausubel, *Current Protocols in Molecular Biology*, John Wiley, 1987-1998, and Sambrook, *et al.*, *Molecular Cloning, A Laboratory Manual*, 2<sup>nd</sup> Edition, Cold Spring Harbor Laboratory Press, 1989.

HCV replicons containing adaptive mutations can be built around an NS3 region or NS5A region containing one or more adaptive mutations described herein. The final replicon will contain replicon components needed for replication and may contain additional components.

SEQ. ID. NO. 2 can be used as a reference point for different HCV regions as follows:

- 5' UTR- nucleotides 1-341;
- Core- nucleotides 342-914;
- E1- nucleotides 915-1490;
- E2- nucleotides 1491-2579;
- P7- nucleotides 2580-2768;
- NS2- nucleotides 2769-3419;
- NS3- nucleotides 3420-5312;
- NS4A- nucleotides 5313-5474;
- NS4B- nucleotides 5475-6257;
- NS5A- nucleotides 6258-7598;
- NS5B- nucleotides 7599-9371; and
- 3' UTR- nucleotides 9374-9605.

The amino acid sequences of the different structural and non-structural regions is provided by SEQ. ID. NO. 1.

- Nucleic acid sequences encoding for a particular amino acid can be produced taking into account the degeneracy of the genetic code. The degeneracy of the genetic code arises because almost all amino acids are encoded for by different combinations of nucleotide triplets or "codons". The translation of a particular codon
- 5 into a particular amino acid is well known in the art (*see, e.g., Lewin GENES IV, p. 119, Oxford University Press, 1990*). Amino acids are encoded for by RNA codons as follows:
- A=Ala=Alanine: codons GCA, GCC, GCG, GCU  
 C=Cys=Cysteine: codons UGC, UGU
- 10 D=Asp=Aspartic acid: codons GAC, GAU  
 E=Glu=Glutamic acid: codons GAA, GAG  
 F=Phe=Phenylalanine: codons UUC, UUU  
 G=Gly=Glycine: codons GGA, GGC, GGG, GGU  
 H=His=Histidine: codons CAC, CAU
- 15 I=Ile=Isoleucine: codons AUA, AUC, AUU  
 K=Lys=Lysine: codons AAA, AAG  
 L=Leu=Leucine: codons UUA, UUG, CUA, CUC, CUG, CUU  
 M=Met=Methionine: codon AUG  
 N=Asn=Asparagine: codons AAC, AAU
- 20 P=Pro=Proline: codons CCA, CCC, CCG, CCU  
 Q=Gln=Glutamine: codons CAA, CAG  
 R=Arg=Arginine: codons AGA, AGG, CGA, CGC, CGG, CGU  
 S=Ser=Serine: codons AGC, AGU, UCA, UCC, UCG, UCU  
 T=Thr=Threonine: codons ACA, ACC, ACG, ACU
- 25 V=Val=Valine: codons GUA, GUC, GUG, GUU  
 W=Trp=Tryptophan: codon UGG  
 Y=Tyr=Tyrosine: codons UAC, UAU.

- Constructs, including subgenomic and genomic replicons, containing one or more of the adaptive mutations described herein can also contain additional
- 30 mutations. The additional mutations may be adaptive mutations and mutations not substantially inhibiting replicon activity. Mutations not substantially inhibiting replicon activity provide for a replicon that can be introduced into a cell and have detectable activity.

### HCV Replicon

HCV replicons include the full length HCV genome and subgenomic constructs. A basic HCV replicon is a subgenomic construct containing an HCV 5' UTR- PC region, an HCV NS3-NS5B polyprotein encoding region, and a HCV 3' UTR. Other nucleic acid regions can be present such as those providing for HCV NS2, structural HCV protein(s) and non-HCV sequences.

The HCV 5' UTR-PC region provides an internal ribosome entry site (IRES) for protein translation and elements needed for replication. The HCV 5'UTR-PC region includes naturally occurring HCV 5' UTR extending about 36 nucleotides into a HCV core encoding region, and functional derivatives thereof. The 5'-UTR-PC region can be present in different locations such as site downstream from a sequence encoding a selection protein, a reporter, protein, or an HCV polyprotein.

Functional derivatives of the 5'-UTR-PC region able to initiate translation and assist replication can be designed taking into structural requirements for HCV translation initiation. (See, for example, Honda, *et al.*, 1996. *Virology* 222, 31-42). The affect of different modifications to a 5' UTR-PC region can be determined using techniques that measure replicon activity.

In addition to the HCV 5' UTR-PC region, non-HCV IRES elements can also be present in the replicon. The non-HCV IRES elements can be present in different locations including immediately upstream the region encoding for an HCV polyprotein. Examples of non-HCV IRES elements that can be used are the EMCV IRES, poliovirus IRES, and bovine viral diarrhea virus IRES.

The HCV 3' UTR assists HCV replication. HCV 3' UTR includes naturally occurring HCV 3' UTR and functional derivatives thereof. Naturally occurring 3' UTR's include a poly U tract and an additional region of about 100 nucleotides. (Tanaka, *et al.*, 1996. *J. Virol.* 70, 3307-3312, Kolykhalov, *et al.*, 1996. *J. Virol.* 70, 3363-3371.) At least *in vivo*, the 3' UTR appears to be essential for replication. (Kolykhalov, *et al.*, 2000. *J. Virol.* 2000 4, 2046-2051.) Examples of naturally occurring 3' UTR derivatives are described by Bartenschlager International Publication Number EP 1 043 399.

The NS3-NS5B polyprotein encoding region provides for a polyprotein that can be processed in a cell into different proteins. Suitable NS3-NS5B polyprotein sequences that may be part of a replicon include those present in different HCV strains and functional equivalents thereof resulting in the processing of NS3-NS5B to

a produce functional replication machinery. Proper processing can be measured for by assaying, for example, NS5B RNA dependent RNA polymerase.

The ability of an NS5B protein to provide RNA polymerase activity can be measured using techniques well known in the art. (See, for example, De  
5 Franscesco, *et al.*, International Publication Number WO 96/37619, Behrens, *et al.*, 1996. *EMBO* 15:12-22, Lohmann, *et al.*, 1998. *Virology* 249:108-118.) Preferably, the sequence of the active NS5B is substantially similar as that provided in SEQ. ID. NO. 1, or a wild type NS5B such as strains HCV-1, HCV-2, HCV-BK, HCV-J, HCV-N, HCV-H. A substantially similar sequence provides detectable HCV polymerase  
10 activity and contains 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acid alterations to that present in a HCV NS5B polymerase. Preferably, no more than 1, 2, 3, 4 or 5 alterations are present.

Alterations to an amino acid sequence provide for substitution(s), insertion(s), deletion(s) or a combination thereof. Sites of different alterations can be  
15 designed taking into account the amino acid sequences of different NS5B polymerases to identify conserved and variable amino acid, and can be empirically determined.

HCV replicons can be produced in a wide variety of different cells and *in vitro*. Suitable cells allow for the transcription of a nucleic acid encoding for an HCV replicon.

20

#### Additional Sequences

An HCV replicon may contain non-HCV sequences in addition to HCV sequences. The additional sequences should not prevent replication and expression, and preferably serve a useful function. Sequences that can be used to  
25 serve a useful function include a selection sequence, a reporter sequence, transcription elements and translation elements.

#### Selection Sequence

A selection sequence in an HCV replicon facilitates the identification  
30 of a cell containing the replicon. Selection sequences are typically used in conjunction with some selective pressure that inhibits growth of cells not containing the selection sequence. Examples of selection sequences include sequences encoding for antibiotic resistance and ribozymes.

Antibiotic resistance can be used in conjunction with an antibiotic to  
35 select for cells containing replicons. Examples of selection sequences providing for

antibiotic resistance are sequences encoding resistance to neomycin, hygromycin, puromycin, or zeocin.

5 A ribozyme serving as a selection sequence can be used in conjunction with an inhibitory nucleic acid molecule that prevents cellular growth. The ribozyme recognizes and cleaves the inhibitory nucleic acid.

#### Reporter Sequence

10 A reporter sequence can be used to detect replicon replication or protein expression. Preferred reporter proteins are enzymatic proteins whose presence can be detected by measuring product produced by the protein. Examples of reporter proteins include, luciferase, beta-lactamase, secretory alkaline phosphatase, beta-glucuronidase, green fluorescent protein and its derivatives. In addition, a reporter nucleic acid sequence can be used to provide a reference sequence that can be targeted by a complementary nucleic acid. Hybridization of the complementary nucleic acid to  
15 its target can be determined using standard techniques.

#### Additional Sequence Configuration

20 Additional non-HCV sequences are preferable 5' or 3' of an HCV replicon genome or subgenomic genome region. However, the additional sequences can be located within an HCV genome as long as the sequences do not prevent detectable replicon activity. If desired, additional sequences can be separated from the replicon by using a ribozyme recognition sequence in conjunction with a ribozyme.

25 Additional sequences can be part of the same cistron as the HCV polyprotein or can be a separate cistron. If part of the same cistron, the selection or reporter sequence coding for a protein should result in a product that is either active as a chimeric protein or is cleaved inside a cell so it is separated from HCV protein.

30 Selection and reporter sequences encoding for a protein when present as a separate cistron should be associated with elements needed for translation. Such elements include a 5' IRES.

#### Detection Methods

35 Methods for detecting replicon activity include those measuring the production or activity of replicon RNA and encoded for protein. Measuring includes qualitative and quantitative analysis.

Techniques suitable for measuring RNA production include those detecting the presence or activity of RNA. The presence of RNA can be detected using, for example, complementary hybridization probes or quantitative PCR. Techniques for measuring hybridization between complementary nucleic acid and  
5 quantitative PCR are well known in the art. (See for example, Ausubel, *Current Protocols in Molecular Biology*, John Wiley, 1987-1998, Sambrook, *et al.*, *Molecular Cloning, A Laboratory Manual*, 2<sup>nd</sup> Edition, Cold Spring Harbor Laboratory Press, 1989, and U.S. Patent No. 5,731,148.)

RNA enzymatic activity can be provided to the replicon by using a  
10 ribozyme sequence. Ribozyme activity can be measured using techniques detecting the ability of the ribozyme to cleave a target sequence.

Techniques for measuring protein production include those detecting the presence or activity of a produced protein. The presence of a particular protein can be determined by, for example, immunological techniques. Protein activity can  
15 be measured based on the activity of an HCV protein or a reporter protein sequence.

Techniques for measuring HCV protein activity vary depending upon the protein that is measured. Techniques for measuring the activity of different non-structural proteins such as NS2/3, NS3, and NS5B, are well known in the art. (See, for example, references provided in the Background of the Invention.)

20 Assays measuring replicon activity also include those detecting virion production from a replicon that produces a virion; and those detecting a cytopathic effect from a replicon producing proteins exerting such an effect. Cytopathic effects can be detected by assays suitable to measure cell viability.

Assays measuring replicon activity can be used to evaluate the ability  
25 of a compound to modulate HCV activities. Such assays can be carried out by providing one or more test compounds to a cell expressing an HCV replicon and measuring the effect of the compound on replicon activity. If a preparation containing more than one compound is found to modulate replicon activity, individual compounds or smaller groups of compounds can be tested to identify replicon active  
30 compounds.

Compounds identified as inhibiting HCV activity can be used to produce replicon enhanced cells and may be therapeutic compounds. The ability of a compound to serve as a therapeutic compound can be confirmed using animal models such as a chimpanzee to measure efficacy and toxicity.

35

### Replicon Enhanced Host Cell

Replicon enhanced cells are initially produced by selecting for a cell able to maintain an HCV replicon and then curing the cell of the replicon. Cells produced in this fashion were found to have an increased ability to maintain a replicon upon subsequent HCV replicon transfection.

Initial transfection can be performed using a wild-type replicon or a replicon containing one or more adaptive mutations. If a wild-type replicon is employed, the replicon should contain a selection sequence to facilitate replicon maintenance.

Cells can be cured of replicons using different techniques such as those employing replicon inhibitory agent. In addition, replication of HCV replicons is substantially reduced in confluent cells. Thus, it is conceivable to cure cells of replicons by culturing them at a high density.

Replicon inhibitory agents inhibit replicon activity or select against a cell containing a replicon. An example of such an agent is IFN- $\alpha$ . Other HCV inhibitory compounds may also be employed. HCV inhibitor compounds are described, for example, in Llinas-Brunet, *et al.*, 2000. *Bioorg Med Chem. Lett.* 10(20), 2267-2270.

The ability of a cured cell to be a replicon enhanced cell can be measured by introducing a replicon into the cell and determining efficiency of subsequent replicon maintenance and activity.

### EXAMPLES

Examples are provided below to further illustrate different features of the present invention. The examples also illustrate useful methodology for practicing the invention. These examples do not limit the claimed invention.

#### Example 1: Techniques

This example illustrates the techniques employed for producing and analyzing adaptive mutations and replicon enhanced cells.

#### *Manipulation of Nucleic Acids and Construction of Recombinant Plasmids*

Manipulation of nucleic acids was done according to standard protocols. (Sambrook, *et al.*, 1989. *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.) Plasmid DNA was

prepared from ON culture in LB broth using Qiagen 500 columns according to manufacturer instructions.

Plasmids containing desired mutations were constructed by restriction digestion using restriction sites flanking the mutations or by PCR amplification of the area of interest, using synthetic oligonucleotides with the appropriate sequence. Site directed mutagenesis was carried out by inserting the mutations in the PCR primers. PCR amplification was performed using high fidelity thermostable polymerases or mixtures of polymerases containing a proofreading enzyme. (Barnes, *et al.*, 1994. *Proc. Natl. Acad. Sci.* 91, 2216-2220.) All plasmids were verified by restriction mapping and sequencing.

pHCVneo17.wt contains the cDNA for an HCV bicistronic replicon identical to replicon I<sub>377</sub>neo/NS3-3'/wt described by Bartenschlager (SEQ. ID. NO. 3) (Lohmann, *et al.*, 1999. *Science* 285,110-113, EMBL-genbank No. AJ242652). The plasmid comprises the following elements: 5' untranslated region of HCV comprising the HCV-IRES and part of the core (nt1-377); neomycin phosphotransferase coding sequence; and EMCV IRES; HCV coding sequences from NS3 to NS5B; 3' UTR of HCV.

Plasmid pHCVNeo17.GAA is identical to pHCVNeo17, except that the GAC triplets (nt. 6934-6939 of pHCVNeo17 sequence) coding for the catalytic aspartates of the NS5B polymerase (amino acids 2737 and 2738 of HCV polyprotein) were changed into GCG, coding for alanine.

Plasmid pHCVNeo17.m0 is identical to pHCVNeo17, except that the triplet AGC (nt. 5335-5337 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2204 of HCV polyprotein) was changed into AGA, coding for arginine.

Plasmid pHCVNeo17.m1 is identical to pHCVNeo17, except that the triplet AAC (nt. 4846-4848 of pHCVNeo17 sequence) coding for the asparagine of NS5A protein (amino acid 2041 of HCV polyprotein) was changed into ACC, coding for threonine.

Plasmid pHCVNeo17.m2 is identical to pHCVNeo17, except that the triplet TCC (nt. 5242-5244 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2173 of HCV polyprotein) was changed into TTC, coding for phenylalanine.



Plasmid pHCVNeo17.m3 is identical to pHCVNeo17, except that the triplet TCC (nt. 5314-5316 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2197 of HCV polyprotein) was changed into TTC, coding for phenylalanine.

5 Plasmid pHCVNeo17.m4 is identical to pHCVNeo17, except that the triplet TTG (nt. 5317-5319 of pHCVNeo17 sequence) coding for the leucine of NS5A protein (amino acid 2198 of HCV polyprotein) was changed into TCG, coding for serine.

10 Plasmid pHCVNeo17.m5 is identical to pHCVNeo17, except that an extra triplet AAA coding for lysine was inserted after the triplet GTG (nt. 4840-4843 of pHCVNeo17 sequence), coding for valine 2039 of HCV polyprotein.

Plasmid pHCVNeo17.m6 is identical to pHCVNeo17, except that the triplets GAA and GCC (nt. 2329-2331 and 2764-2766 of pHCVNeo17 sequence) coding for the glutamic acid and the alanine of NS3 protein (amino acid 1202 and 1347 of HCV polyprotein) were changed respectively into GGA and ACC, coding for glycine and threonine. The triplet TCC (nt. 5242-5244 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2173 of HCV polyprotein) was changed into TTC, coding for phenylalanine; an extra adenosine was inserted into the EMCV IRES (after the thymidine 1736 of the replicon sequence).

20 Plasmid pHCVNeo17.m7 is identical to pHCVNeo17, except that the triplet AAC (nt. 4846-4848 of pHCVNeo17 sequence) coding for the asparagine of NS5A protein (amino acid 2041 of HCV polyprotein) was changed into ACC, coding for threonine; the triplet TCC (nt. 5242-5244 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2173 of HCV polyprotein) was changed into 25 TTC, coding for phenylalanine.

Plasmid pHCVNeo17.m8 is identical to pHCVNeo17, except that the triplet AAC (nt. 4846-4848 of pHCVNeo17 sequence) coding for the asparagine of NS5A protein (amino acid 2041 of HCV polyprotein) was changed into ACC, coding for threonine; the triplet TCC (nt. 5314-5316 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2197 of HCV polyprotein) was changed into 30 TTC, coding for phenylalanine.

Plasmid pHCVNeo17.m9 is identical to pHCVNeo17, except that the triplet AAC (nt. 4846-4848 of pHCVNeo17 sequence) coding for the asparagine of NS5A protein (amino acid 2041 of HCV polyprotein) was changed into ACC, coding

for threonine; the triplet TTG (nt. 5317-5319 of pHCVNeo17 sequence) coding for the leucine of NS5A protein (amino acid 2198 of HCV polyprotein) was changed into TCG, coding for serine.

5 Plasmid pHCVNeo17.m10 is identical to pHCVNeo17, except that the triplet GAA (nt. 2329-2331 of pHCVNeo17 sequence) coding for the glutamic acid of NS3 protein (amino acid 1202 of HCV polyprotein) was changed into GGA, coding for glycine; an extra triplet AAA coding for lysine was inserted after the triplet GTG (nt. 4840-4843 of pHCVNeo17 sequence), coding for valine 2039 of HCV polyprotein.

10 Plasmid pHCVNeo17.m11 is identical to pHCVNeo17, except that the triplet TCC (nt. 5314-5316 of pHCVNeo17 sequence) coding for the serine of NS5A protein (amino acid 2197 of HCV polyprotein) was changed into TTC, coding for phenylalanine. The triplet GCC (nt. 5320-5322 of pHCVNeo17 sequence) coding for the alanine of NS5A protein (amino acid 2199 of HCV polyprotein) was changed into  
15 ACC coding for threonine.

Plasmid pHCVNeo17.m12 is identical to pHCVNeo17, except that the triplet AAC (nt. 4846-4848 of pHCVNeo17 sequence) coding for the asparagine of NS5A protein (amino acid 2041 of HCV polyprotein) was changed into ACC, coding for threonine; the triplet TCC (nt. 5314-5316 of pHCVNeo17 sequence) coding for  
20 the serine of NS5A protein (amino acid 2197 of HCV polyprotein) was changed into TTC, coding for phenylalanine. The triplet GCC (nt. 5320-5322 of pHCVNeo17 sequence) coding for the alanine of NS5A protein (amino acid 2199 of HCV polyprotein) was changed into ACC coding for threonine.

Plasmid pHCVNeo17.m13 has the same mutations as  
25 pHCVNeo17.m8, but also an extra adenosine inserted into the EMCV IRES (after the thymidine 1736 of the replicon sequence).

Plasmid pHCVNeo17.m14 has the same mutations as pHCVNeo17.m11, but also an extra adenosine inserted into the EMCV IRES (after the thymidine 1736 of the replicon sequence).

30 Plasmid pHCVNeo17.m15 is identical to pHCVNeo17, except that the triplet GCC (nt. 5320-5322 of pHCVNeo17 sequence) coding for the alanine of NS5A protein (amino acid 2199 of HCV polyprotein) was changed into ACC coding for threonine.

Plasmid pRBSEAP.5 is a pHCVNeo.17 derivative where the Neo coding sequence has been replaced with the sequence coding for the human placental alkaline phosphatase corresponding to nucleotides 90-1580 of pBC12/RSV/SEAP plasmid. (Berger, *et al.*, 1988. *Gene* 66, 1-10.)

5

#### *RNA Transfection*

Transfection was performed using Huh-7 cells. The cells were grown in Dulbecco's modified minimal essential medium (DMEM, Gibco, BRL) supplemented with 10% FCS. For routine work, cells were passed 1 to 5 twice a week using 1x trypsin/EDTA (Gibco, BRL).

10

Plasmids were digested with the ScaI endonuclease (New England Biolabs) and transcribed *in vitro* with the T7 Megascript kit (Ambion). Transcription mixtures were treated with DNase I (0.1 U/ml) for 30 minutes at 37°C to completely remove template DNA, extracted according to the procedure of Chomczynski (Chomczynski, *et al.*, 1987. *Anal. Biochem.* 162, 156-159), and resuspended with RNase-free phosphate buffered saline (rfPBS, Sambrook, *et al.*, 1989. *Molecular Cloning: A Laboratory Manual*, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).

15

RNA transfection was performed as described by Liljestrom, *et al.*, 1991. *J. Virol.* 6, 4107-4113, with minor modifications. Subconfluent, actively growing cells were detached from the tissue culture container using trypsin/EDTA. Trypsin was neutralised by addition of 3 to 10 volumes of DMEM/10%FCS and cells were centrifuged for 5 minutes at 1200 rpm in a Haereus table top centrifuge at 4°C. Cells were resuspended with ice cold rfPBS by gentle pipetting, counted with a haemocytometer, and centrifuged as above. rfPBS wash was repeated once and cells were resuspended at a concentration of  $1-2 \times 10^7$  cell/ml in rfPBS. Aliquots of cell suspension were mixed with RNA in sterile eppendorf tubes. The RNA/cell mixture was immediately transferred into the electroporation cuvette (precooled on ice) and pulsed twice with a gene pulser apparatus equipped with pulse controller (Biorad). Depending on the experiment, 0.1, 0.2 or 0.4 cm electrode gap cuvettes were used, and settings adjusted (Table 3).

20

25

30

TABLE 3

| Cuvette  | Volume     | Voltage | Capacitance | Resistance | RNA        |
|----------|------------|---------|-------------|------------|------------|
| gap (cm) | ( $\mu$ l) | (Volts) | ( $\mu$ Fa) | (ohm)      | ( $\mu$ g) |
| 0.1      | 70         | 200     | 25          | infinite   | 1-10       |
| 0.2      | 200        | 400     | 25          | infinite   | 5-20       |
| 0.4      | 800        | 800     | 25          | infinite   | 15-100     |

After the electric shock, cells were left at room temperature for 1-10 minutes (essentially the time required to electroporate all samples) and subsequently diluted with at least 20 volumes of DMEM/10%FCS and plated as required for the experiment. Survival and transfection efficiency were monitored by measuring the neutral red uptake of cell cultured for various days in the absence or in the presence of neomycin sulfate (G418). With these parameters, survival of Huh-7 cells was usually 40-60% and transfection efficiency ranged between 40% and 100%.

#### *Sequence Analysis of Replicon RNAs*

The entire NS region was recloned from 3 different transfection experiments performed with HCVNeo.17 RNA. RNA was extracted from selected clones either using the Qiagen RNAeasy minikit following manufacturer instructions or as described by Chomczynski, *et al.*, 1987. Anal. Biochem. 162, 156-159.

Replicon RNAs (5  $\mu$ g of total cellular RNA) were retro-transcribed using oligonucleotide HCVG34 (5'- ACATGATCTGCAGAGAGGCCAGT-3'; SEQ. ID. No. 4) and the Superscript II reverse transcriptase (Gibco, BRL) according to manufacturer instructions, and subsequently digested with 2 U/ml Ribonuclease H (Gibco BRL). The cDNA regions spanning from the EMCV IRES to the HCV 3' end were amplified by PCR using oligonucleotides HCVG39 (5'- GACASGCTGTGATAWATGTCTCCCCC-3'; SEQ. ID. NO. 5) and CITE3 (5'- TGGCTCTCCTCAAGCGTATTC -3'; SEQ. ID. NO. 6) and the LA Taq DNA polymerase (Takara LA Taq).

Amplified cDNAs were digested with the KpnI endonuclease (New England Biolabs) and the 5.8 kb fragments were gel purified and ligated to the 5.6 kb vector fragment (purified from plasmid pRBSEAP.5 digested with KpnI) using T4

DNA ligase (New England Biolabs) according to manufacturer instructions. Ligated DNAs were transformed by electroporation in DH10B or JM119 strains of *E. coli*.

In the case of NS5A region, total RNA isolated from 3 clones, (HB77, HB60 and HB68) was extracted and used for RT-PCR. 5µg of total RNA plus 20 pmole of AS61 oligo (5'-ACTCTCTGCAGTCAAGCGGCTCA-3', RT antisense oligo; SEQ. ID. NO. 7) were heated 5 minutes at 95°C, then DMSO (5% f.c.), DTT (10 mM f.c.), 1 mM dNTP (1 mM f.c.), 1x Superscript buffer (1 x f.c.), and 10 u Superscript (Gibco) were added to a total volume of 20 µl and incubated 3 hours at 42°C. 2µl of this RT reaction were used to perform PCR with oligos S39 (5'-CAGTGGATGAACCGGCTGATA-3', sense; SEQ. ID. NO. 8) or S41 (5'-GGGGCGACGGCATCATGCAAACC-3', sense; SEQ. ID. NO. 9) and B43 (5'-CAGGACCTGCAGTCTGTCAAAGG-3', antisense; SEQ. ID. NO. 10) using Elongase Enzyme Mix (Gibco) according the instruction provided by the manufacturer. The resulting PCR fragment was cloned in pCR2.1 vector using the TA Cloning kit (Invitrogen) and transformed in Top10F' bacterial strain.

Plasmid DNA was prepared from ON culture of the resulting ampicillin resistant colonies using Qiagen 500 columns according to manufacturer instructions. The presence of the desired DNA insert was ascertained by restriction digestion, and the nucleotide sequence of each plasmid was determined by automated sequencing. Nucleotide sequences and deduced amino acids sequences were aligned using the GCG software.

#### *TaqMan*

TaqMan analysis was typically performed using 10 ng of RNA in a reaction mix (TaqMan Gold RT-PCR kit, Perkin Elmer Biosystems) either with HCV specific oligos/probe (oligo 1: 5'-CGGGAGAGCCATAGTGG-3'; SEQ. ID. NO. 11, oligo 2: 5'-AGTACCACAAGGCCTTTCG-3'; SEQ. ID. NO. 12, probe: 5'-CTGCGGAACCGGTGAGTACAC-3'; SEQ. ID. NO. 13) or with human GAPDH specific oligos/probe (Pre-Developed TaqMan Assay Reagents, Endogenous Control Human GAPDH, Part Number 4310884E, Perkin Elmer Biosystems). PCR was performed using a Perkin Elmer ABI PRISM 7700 under the following conditions: 30 minutes at 48°C (the RT step), 10 minutes at 95°C and 40 cycles: 15 seconds at 95°C and 1 minute at 60°C. Quantitative calculations were obtained using the Comparative C<sub>T</sub> Method (described in User Bulletin #2, ABI PRISM 7700 Sequence Detection System, Applied Biosystem, Dec 1997) considering the level of GAPDH mRNA

constant. All calculations of HCV RNA are expressed as fold difference over a specific control.

#### *Antibodies and Immunological Techniques*

5                    Mouse monoclonal antibody (anti-NS3 mab10E5/24) were produced by standard techniques. (Galfré and Milstein, 1981. *Methods in Enzymology* 73, 1-46.) Purified recombinant protein was used as an immunogen. (Gallinari, *et al.*, 1999. *Biochemistry* 38, 5620-5632.)

10                    For Cell-ELISA analysis, transfected cells were monitored for expression of the NS3 protein by ELISA with the anti-NS3 mab 10E5/24. Cells were seeded into 96 well plates at densities of 40,000, 30,000, 15,000 and 10,000 cells per well and fixed with ice-cold isopropanol at 1, 2, 3 and 4 days post-transfection, respectively. The cells were washed twice with PBS, blocked with 5% non-fat dry milk in PBS + 0.1% Triton X100 + 0.02% SDS (PBSTS) and then incubated  
15                    overnight at 4°C with 10E5/24 mab diluted 1:2000 in Milk/PBSTS. After washing 5 times with PBSTS, the cells were incubated for 3 hours at room temperature with anti-mouse IgG Fc specific alkaline phosphatase conjugated secondary antibody (Sigma A-7434), diluted 1:2000 in Milk/PBSTS. After washing again as above, the reaction was developed with *p*-nitrophenyl phosphate disodium substrate (Sigma 104-  
20                    105) and the absorbance at 405 nm read at intervals.

                      The results were normalized by staining with sulforhodamine B (SRB Sigma S 1402) to determine cell numbers. The alkaline phosphatase substrate was removed from the wells and the cells washed with PBS. The plates were then incubated with 0.4% SRB in 1% acetic acid for 30 minutes (200 µl/well), rinsed 4  
25                    times in 1% acetic acid, blotted dry and then 200 µl/well of 10mM Tris pH 10.5 added. After mixing, the absorbance at 570 nm was read.

#### *Neutral Red/ Crystal Violet Staining of Foci*

                      The survival of transfected cells in the absence or presence of G418  
30                    was monitored by staining of foci/clones with neutral red *in vivo* with subsequent crystal violet staining. The medium was removed from the cells and replaced with fresh medium containing 0.0025% neutral red (Sigma N2889) and the cells incubated for 3 hours at 37°C. Cells were washed twice with PBS, fixed in 3.5% formaldehyde for 15 minutes, washed twice again in PBS and then with distilled water and the  
35                    number of (live) foci counted. The cells could then be re-stained with crystal violet

by incubating with an 0.1% crystal violet (Sigma C0775) solution in 20% methanol for 20 minutes at room temperature, followed by 3 washes in 20% methanol and a wash with distilled water.

#### 5 *Preparation Of Cells Cured Of Endogenous Replicon*

Replicon enhanced cells designated 10IFN and Cl.60/cu were produced using different HCV inhibitory agents. Based on the techniques described herein additional replicon enhanced clones can readily be obtained.

- 10IFN was obtained by curing a Huh-7 cell of a replicon using human
- 10 IFN- $\alpha$ 2b. Huh-7 cells containing HCV replicons (designated HBI10, HBIII4, HBIII27 and HBIII18) were cultured for 11 days in the presence of 100 U/ml recombinant human IFN- $\alpha$ 2b (Intron-A, Schering-Plough), and subsequently for 4 days in the absence of IFN- $\alpha$ 2b. At several time points during this period, the clones were analyzed for the presence of HCV proteins and RNA by Western and Northern
- 15 blotting. After 7 days of incubation with IFN- $\alpha$ 2b, HCV proteins could no longer be detected in any of these clones by Western blotting and similar effects were seen with RNA signals in Northern blots. IFN- $\alpha$ 2b treated cells were stored in liquid nitrogen until used for transfection experiments.

- Cl.60/cu was obtained by curing a Huh-7 cell of a replicon using an
- 20 HCV inhibitory compound. The presence of HCV RNA was determined using PCR (TaqMan) at 4, 9, 12 and 15 days. From day 9 the amount of HCV RNA was below the limit of detection. To further test the disappearance of the replicon, 4 million cells of cured Clone 60 cells (after the 15 days of treatment) were plated in the presence of 1 mg/ml G-418. No viable cells were observed, confirming that absence of HCV
- 25 replicons able to confer G-418 resistance.

#### Example 2: Selection and Characterization of Cell Clones Containing Functional HCV Replicons

- Huh-7 cells ( $2-8 \times 10^6$ ) were transfected by electroporation with *in vitro*
- 30 transcribed replicon RNAs (10–20  $\mu$ g), plated at a density ranging from  $2.5 \times 10^3$  to  $10 \times 10^3/\text{cm}^2$ , and cultured in the presence of 0.8–1 mg/ml G418. The majority of replicon transfected cells became transiently resistant to G418 and duplicated normally for 7 to 12 days in the presence of the drug, while cells transfected with irrelevant RNAs and mock transfected cells did not survive more than 7 days (data not
- 35 shown). Transient resistance to G418 was likely due to persistence of the Neo protein

expressed from the transfected RNA, since it was observed also with mutated replicons unable to replicate. Approximately 2 weeks after transfection, transient resistance declined, most cells died and small colonies of cells permanently resistant to the antibiotic became visible in samples transfected with HCVNeo.17 RNA, but not in cells transfected with other replicon RNAs.

In several experiments, the frequency of G418 resistant clones ranged between 10 and 100 clones per  $10^6$  transfected cells. About 20 G418 resistant colonies were isolated, expanded and molecularly characterized. PCR and RT-PCR analysis of nucleic acids indicated that all clones contained HCV RNA but not HCV DNA, demonstrating that G418 resistance was due to the presence of functional replicons (data not shown). This result was confirmed by Northern blot analysis and metabolic labeling with  $^3\text{H}$ -uridine, which revealed the presence of both genomic and antigenomic HCV RNAs of the expected size (data not shown). Lastly, western blot, immunoprecipitation and immunofluorescence experiments showed that these clones expressed all HCV non-structural proteins as well as Neo protein (data not shown).

Clones differed in terms of cell morphology and growth rate. Replicon RNA copy number (500-10000 molecules/cell) and viral protein expression also varied between different clones (data not shown). However, the amount of replicon RNA and proteins also varied with passages and with culture conditions and was higher when cells were not allowed to reach confluency, suggesting that replicons replicated more efficiently in dividing cells than in resting cells. Processing of the viral polyprotein occurred with kinetics similar to those observed in transfected cells.

Interestingly, in all tested clones HCV replication was efficiently inhibited by treating the cells with IFN- $\alpha$ 2b. The  $\text{EC}_{50}$  was between 1 and 10 U/ml, depending on the clone.

### Example 3: Identification of Adaptive Mutations

The low number of G418 resistant clones derived from HCVNeo.17 RNA transfection suggested that replication could require mutation(s) capable of adapting the replicon to the host cell (adaptive mutations) and/or that only a small percentage of Huh-7 cells were competent for HCV replication. To verify the first hypothesis, mutations in replicons RNAs derived from selected cell clones were identified.



RNA sequences for different replicons were determined using standard techniques. Such techniques involved isolating RNA from several independent clones, reverse transcription to produce cDNA, amplifying cDNAs by PCR and cloning into an appropriate vector. The cDNA spanning almost the entire HCV NS region (126 bp at the 3' end of the EMCV IRES and 5650 bp of the HCV NS region (i.e., the entire NS ORF and 298 nucleotides at the 3' end) from 5 clones (HBI10, HBIII12, HBIII18, HBIII27, HBIV1) were recloned and sequenced. In addition, the NS5A coding region (nt. 4784-6162) from 3 additional clones (HB 77, HB 68 and HB 60) were recloned and sequenced.

To discriminate mutations present in the replicon RNA from those derived from the cloning procedure, at least 2 isolates derived from independent RT-PCR experiments were sequenced for each cell clone. Comparison of the nucleotide sequences with the parental sequence indicated that each isolate contained several mutations (Tables 4A and 4B).

TABLE 4A

| Cell clone       | HBIII 12   |  | HBIII 18   |  | HBI 10                                       |                                  | HBIII 27   |  |
|------------------|--|--|--|--|--|----------------------------------|--|--|
| isolate          | 4  | 29   | 28   | 61   | 12   | 43                               | 13   | 72   |
|                  | 1674-7460  | 1674-7460                                    | 1674-7460  | 1674-7460  | 1674-7460                                    | 1674-7460                        | 1674-7460  | 1674-7460  |
| EMCV IRES 126 bp | A @ 1736   | A @ 1736                                     |  | C 1752 T   |  |                                  |  | T 1678 C   |
| NS3 1895 bp      | G 2009 C<br>A 2698 G<br>G 2764 A<br>A 3256 G<br>T 3273 C | A 2330 G<br>C 2505 T<br>G 2764 A<br>T 3085 C | T 2150 C<br>C 2196 A<br>T 3023 A<br>T 3134 C<br>C 3267 T | T 2015 C<br>A 2338 G<br>C 2616 T<br>A 2664 G<br>A 3148 G<br>T 3286 C<br>C 3615 T<br>C 3657 T | T 1811 A<br>A 2330 G<br>T 2666 C<br>T 3395 C | A 2330 G<br>A 2882 G<br>T 3673 C | G 2009 C<br>T 2015 C<br>C 2336 G<br>A 3130 T<br>A 3401 G<br>A 3518 C | G 2009 C<br>C 2052 A<br>G 2644 A<br>C 2803 A<br>T 2823 A<br>T 3692 C |
| NS4A 161 bp      | T 3790 C   |  | A 3847 G   | T 3827 A   | T 3742 C                                     |                                  | A 3743 G   | A 3797 G   |
| NS4B 782 bp      | T 3869 C<br>A 4107 G<br>T 4185 C<br>A 4428 G             | C 4283 T<br>C 4429 T                         | G 4300 A   | A 4136 G<br>A 4261 G<br>G 4309 A<br>A 4449 G   | T 4290 C                                     | A 4053 G<br>A 2496 C<br>T 4316 G | G 3880 A<br>T 4200 C<br>A 4366 G                                     | C 4547 T   |

TABLE 4A

| Cell clone      | HBIII 12   |  | HBIII 18   |  | HBI 10                                |  | HBIII 27   |  |
|-----------------|--|--|--|--|---------------------------------------|--|--|--|
| isolate         | 4  | 29   | 28   | 61   | 12                                    | 43   | 13   | 72   |
|                 | <b>1674-7460</b>   | <b>1674-7460</b>   | <b>1674-7460</b>   | <b>1674-7460</b>   | <b>1674-7460</b>                      | <b>1674-7460</b>   | <b>1674-7460</b>   | <b>1674-7460</b>   |
| NS5A<br>1340 bp | A 4847 C<br>G 5158 A<br><br>G 5175 C<br><br>C 5243 T<br>C 5390 T<br>A 5719 G | G 4728 A<br>A 4845 G<br><br>C 5243 T<br><br>G 5512 T<br>A 5521 G<br>A 5600 G<br>A 5740 C | C 5243 T<br>A 5486 G<br><br>C 5596 T<br><br>G 5823 A     | C 4729 A<br>T 4993 C<br><br>G 5095 A<br><br>T 5334 C<br>A 5374 T<br>T 5379 A<br>T 5480 C<br>A 5513 G<br>T 5977 C | A 4694 T<br>AAA @<br>4842<br>T 5237 C | A 4675 G<br>A 4761 G<br><br>AAA @<br>4842<br>T 5368 C                | A 4855 G<br>C 5006 T<br><br>T 5318 C<br><br>A 5574 G<br>G 5866 A | A 4888 G<br>C 4985 T<br><br>T 5030 A<br><br>T 5090 A<br>T 5318 C<br>A 5328 G<br>A 5399 G<br>A 5574 G |
| NS5B<br>1477 bp | T 6316 C<br>T 6589 C<br>T 7370 C   | A 6406 G<br>G 6756 A<br>G 6963 T   | T 6074 C<br>A 6541 G<br>A 6732 G<br>A 7350 T<br>A 7359 G | A 6150 G<br>A 6218 G<br>T 7352 A   | A 6911 G                              | A 5986 G<br>T 6099 C<br>C 6141 T<br>G 6463 A<br>C 6849 T<br>T 6865 C | G 6479 C<br>C 6870 T<br>A 7213 G<br>T 7448 C                     | G 6156 A<br>G 7434 A<br>T 7444 C   |

Clone name and isolate number are indicated in the first and second row, respectively.

The first and the last nucleotide of the region that was recloned and sequenced are indicated in the third row.

5

Nucleotide (IUB code) substitutions are indicated with the original nucleotide, its position and mutated nucleotide.

Nucleotide(s) insertions are indicated with the nucleotide(s), the symbol @ and the position of the nucleotide preceding insertion.

10

Numbering refers to the first nucleotide of the replicon sequence (EMBL-genbank No. AJ242652).

The region in which mutations are located and the nucleotide length of each region are indicated in the left most column.

Silent mutations are in italic.

Non sense mutations are underlined.

15

Consensus mutations are bold.

TABLE 4B

| Cell clone             | HBIV1            |                      | HB 77            |                  | HB 68            |                  | HB 60            |                  |
|------------------------|------------------|----------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| isolate                | 85               | 93                   | 10               | 14               | 42               | 1                | 13               | 7                |
|                        | <b>1674-7460</b> | <b>1674-7460</b>     | <b>4784-6162</b> | <b>4465-6162</b> | <b>4784-6162</b> | <b>4465-6162</b> | <b>4784-6162</b> | <b>4784-6162</b> |
| EMCV<br>IRES<br>126 bp |                  | A @ 1736             |                  |                  |                  |                  |                  |                  |
| NS3<br>1895 bp         | A 3403 G         | A 2572 G<br>A 3454 G |                  |                  |                  |                  |                  |                  |

TABLE 4B

| Cell clone      | HBIV1  |  | HB 77                            |  | HB 68  |  | HB 60  |                                  |
|-----------------|--|--|----------------------------------|--|--|--|--|----------------------------------|
| isolate         | 85   | 93   | 10                               | 14   | 42   | 1  | 13   | 7                                |
|                 | 1674-7460  | 1674-7460  | 4784-6162                        | 4465-6162  | 4784-6162  | 4465-6162  | 4784-6162  | 4784-6162                        |
| NS4A<br>161 bp  |  |  |                                  |  |  |  |  |                                  |
| NS4B<br>782 bp  | A 4084 G   | C 3892 T   |                                  |  |  |  |  |                                  |
| NS5A<br>1340 bp | T 4742 C<br>C 5315 T<br>G 5431 T<br>T 5751 C<br>T 5797 C   | A 4847 C<br>A 5225 G<br>C 5315 T<br>G 5320 A<br>T 5356 A<br>G 5523 A<br>T 5888 A | C 4813 T<br>G 5060 C<br>C 5337 A | A 4699 C<br>A 5161 G<br>C 5337 A<br>A 5459 G<br>T 5977 C | T 5171 G<br>C 5298 T<br>C 5337 A<br>A 5639 G<br>A 5969 G | T 4587 C<br>T 4972 C<br>A 5094 G<br>A 5278 G<br>G 5320 A<br>C 5532 T | A 4821 G<br>G 5320 A<br>A 5414 G<br>T 5601 G<br>C 5808 T | C 5337 G<br>C 5551 T<br>G 5806 A |
| NS5B<br>1477 bp | T 6144 A<br>A 6365 G<br>A 6656 G<br>A 6677 G<br>T 6855 C<br>T 6947 A<br>T 6997 C<br>G 7041 T<br>A 7187 C | T 6855 C<br>A 7135 G<br>T 7171 C   |                                  |  |  |  |  |                                  |

See Table 4A legend.

- 5 The frequency of mutations ranged between  $1.7 \times 10^{-3}$  and  $4.5 \times 10^{-3}$  (average  $3 \times 10^{-3}$ ). The majority of mutations were nucleotide substitutions, although insertions of 1 or more nucleotides were also observed (Tables 4A and 4B).

- 10 Approximately 85% of the mutations found only in 1 isolate (non-consensus) were randomly distributed in the recloned fragment, and possibly include mis-incorporation during the PCR amplifications. Conversely, the remaining 15% of the mutations were common to 2 or more isolates derived from independent RT-PCR experiments (consensus mutations), and presumably reflected mutations present in the template RNA.

- 15 Consensus mutations were found in all isolates and were either common to isolates derived from the same clone (consensus A), or to isolates derived from different clones (consensus B). Analysis of additional isolates derived from the same cell clones indicated that consensus A mutations were not always present in all isolates derived from one clone (data not shown). This observation, together with the

presence of consensus B mutations, suggests that, even within a single cell clone, replicons exist as quasi-species of molecules with different sequences.

At variance with non-consensus mutations, consensus mutations were not randomly distributed but were clustered in the regions coding for the NS5A protein (frequency  $1 \times 10^{-3}$ ) and for the NS3 protein (frequency  $0.5 \times 10^{-3}$ ). Only one consensus mutation was found in the region coding for the NS5B protein (frequency  $0.1 \times 10^{-3}$  nucleotides) and none in the regions coding for NS4A and NS4B. Interestingly, 1 consensus mutation was observed also in the EMCV IRES.

With the exception of 2 silent mutations found in NS5A and NS5B, consensus mutations occurring in the NS region resulted in changes in the deduced amino acid sequence (Tables 5A and 5B). Noticeably, these amino acid changes occurred in residues that are conserved in all or most natural HCV isolates. Interestingly, clones HB 77 and HB 60 displayed different nucleotide substitutions (C5337A and C5337G, respectively) resulting in the same amino acidic mutation (S 2204 R).

TABLE 5A

| Cell clone | HBIII 12             |                      | HBIII 18 |        | HBI 10   |          | HBIII 27             |                      |
|------------|----------------------|----------------------|----------|--------|----------|----------|----------------------|----------------------|
| isolate    | 4                    | 29                   | 28       | 61     | 12       | 43       | 13                   | 72                   |
| NS3        | G 1095 A<br>A 1347 T | E 1202 G<br>A 1347 T |          |        | E 1202 G | E 1202 G | G 1095 A             | G 1095 A             |
| NS4A       |                      |                      |          |        |          |          |                      |                      |
| NS4B       |                      |                      |          |        |          |          |                      |                      |
| NS5A       | N 2041 T<br>S 2173 F | S 2173 F             | S 2173 F | E 2263 | K @ 2039 | K @ 2039 | L 2198 S<br>R 2283 R | L 2198 S<br>R 2283 R |
| NS5B       |                      |                      |          |        |          |          |                      |                      |

See Table 4A legend.

**TABLE 5B**

| Cell clone | HBIV1    |                                  | HB 77    |          | HB 68    |          | HB 60    |          |
|------------|----------|----------------------------------|----------|----------|----------|----------|----------|----------|
| isolate    | 85       | 93                               | 10       | 14       | 42       | 1        | 13       | 7        |
| NS3        |          |                                  |          |          |          |          |          |          |
| NS4A       |          |                                  |          |          |          |          |          |          |
| NS4B       |          |                                  |          |          |          |          |          |          |
| NS5A       | S 2197 F | N 2041 T<br>S 2197 F<br>A 2199 T | S 2204 R | S 2204 R | S 2204 R | A 2199 T | A 2199 T | S 2204 R |
| NS5B       | N 2710 N | N 2710 N                         |          |          |          |          |          |          |

See Table 4A legend.

#### 5 Example 4: Functional Characterization of Consensus Mutations

The identification of consensus mutations in recloned replicons indicated that replication proficiency of replicon RNAs contained in selected cell clones depended from the presence of such mutations. To substantiate this hypothesis, the effect of several consensus mutations on replication were analyzed.

10 Consensus mutations found in the NS5A region were more closely analyzed. Consensus mutations were segregated from the non-consensus ones, and pHCVNeo.17 derivatives containing single or multiple consensus mutations were constructed (Table 6).

**TABLE 6**

| Construct     | Consensus mutations |        |           | G418 cfu/10 <sup>3</sup><br>transfected<br>cells |
|---------------|---------------------|--------|-----------|--|
|               | NS3                 | NS5A   | EMCV IRES |  |
| pHCVNeo17.wt  |                     |        |           | 0-3  |
| pHCVNeo17.GAA |                     |        |           | 0  |
| pHCVNeo17.m0  |                     | S2204R |           | 30-130   |
| pHCVNeo17.m1  |                     | N2041T |           | 0-3  |
| pHCVNeo17.m2  |                     | S2173F |           | 15-60  |
| pHCVNeo17.m3  |                     | S2197F |           | 160-500  |
| pHCVNeo17.m4  |                     | L2198S |           | 30-50  |

**TABLE 6**

| Construct     | Consensus mutations |                        |           | G418 cfu/10 <sup>5</sup> transfected cells |
|---------------|---------------------|------------------------|-----------|--|
|               | NS3                 | NS5A                   | EMCV IRES |  |
| pHCVNeo17.m5  |                     | <u>K@2039</u>          |           | 25-55                                      |
| pHCVNeo17.m6  | E1202G; A1347T      | S2173F                 | Extra A   | 13-100                                     |
| pHCVNeo17.m7  |                     | N2041T; S2173F         |           | 0-1  |
| pHCVNeo17.m8  |                     | N2041T; S2197F         |           | 360-500                                    |
| pHCVNeo17.m9  |                     | N2041T; L2198S         |           | 140-170                                    |
| pHCVNeo17.m10 | E1202G              | <u>K@2039</u>          |           | 1060                                       |
| pHCVNeo17.m11 |                     | S2197F; A2199T         |           | 900  |
| pHCVNeo17.m12 |                     | N2041T; S2197F; A2199T |           | >1000                                      |
| pHCVNeo17.m13 |                     | N2041T; S2197F         | Extra A   | 100  |
| pHCVNeo17.m14 |                     | S2197F; A2199T         | Extra A   | >500                                       |
| pHCVNeo17.m15 |                     | A2199T                 |           | 300-600                                    |

Huh-7 cells ( $2 \times 10^6$ ) were transfected with 10  $\mu$ g of RNA transcribed from the indicated constructs. Approximately  $2 \times 10^5$  cells were plated in a 10 cm tissue culture dish and cultured with 1 mg/ml G418 for 20 days.

Colonies surviving selection were stained with crystal violet and counted.

5

RNAs transcribed *in vitro* from these constructs were transfected in Huh-7 cells and the affect on replication was estimated by counting neomycin resistant colonies (G418 cfu). As shown in Table 6, all but 1 construct containing single consensus mutations showed a significant increase on G418 cfu efficiency, thus indicating that the corresponding mutations improved replication. Noticeably, 2 mutants containing single mutations in NS5A (m3 and m15) were clearly more effective than all other single mutants. Results of mutants containing 2 or more mutations, indicated the presence of a synergistic effect in some combinations (m8, m9, m11 and possibly m10), but also a slightly antagonistic effect in 1 mutant (m7).

15

#### Example 5: Replicon Replication in the Absence of Selection

Replication of HCV replicons in the absence of a G418 selection was detected using quantitative PCR (TaqMan). At 24 hours post-transfection a large amount of replicon RNA was detected in cells transfected with all replicons, including the GAA control replicon containing mutations in the catalytic GDD motif of the NS5B polymerase. This result suggested that analysis at very early time points (up to 48 hour post-transfection) essentially measured the input RNA. Northern blot analysis also indicated that after 24 hours the majority of the transfected RNA was degraded intracellularly (data not shown).

20

Analysis at later time points showed that the amount of replicon RNA was considerably reduced at 4 days and eventually became undetectable (6/8 days) in cells transfected with replicon HCVNeo17.wt, but was still high in cells transfected with replicons m0, m3 and m15 (Table 7). At day six, that the amount of replicon RNA became undetectable in cells transfected with replicon HCVNeo17.wt, m0, and m2, but was detectable in cells transfected with replicon m3 and m15 (Table 7).

**TABLE 7**

| Name         | Hu H7    |          |
|--------------|----------|----------|
|              | RNA equ. | RNA equ. |
|              | day 4    | day 6    |
| Wt           | 1 x      | 1 x      |
| hcvneo17.m0  | 3 x      | 1 x      |
| hcvneo17.m2  | 1 x      | 1 x      |
| hcvneo17.m3  | 5 x      | 3 x      |
| hcvneo17.m15 | 6 x      | 5 x      |

10

Persistence of m0, m3 and m15 replicons RNA was abolished by treatment with interferon- $\alpha$  or with an HCV inhibitory compound (data not shown). Moreover, RNA persistence was not observed with mutated replicons carrying the NS5B GAA mutation besides adaptive mutations (data not shown). Taken together, these results demonstrated that quantitative PCR could be used to monitor replication at early times post-transfection, and can be used to evaluate the replication proficiency of replicon RNAs containing mutations.

Comparison of the results shown in Tables 6 and 7, indicated that there was a good correlation between the amount of replicon RNA detected by TaqMan and the G418 cfu efficiency. Nonetheless, some mutants (m2, m3) showed a pronounced effect on G418 cfu efficiency, and little if any effect on early replication as measured by TaqMan PCR, while other mutants (m0) showed the reverse behavior.

20

**Example 6: HCV Replicon Enhanced Cells**

HCV replicon enhanced cells were produced by introducing an HCV replicon into a host, then curing the host of the replicon. Adaptive mutations (or combinations of them) by themselves increased up to 2 orders of magnitude the G418 cfu efficiency and enhanced early replication comparably. Nonetheless, even with the most effective mutants, only a small percentage of transfected cells (<5 %, data not shown) gave rise to G418 resistant clones containing functional replicons. This observation was attributed, at least in part to a low cloning efficiency of Huh-7 cells (data not shown), and only a fraction of Huh-7 cells being competent for replication.

Several clones were cured of endogenous replicons by treating them for about 2 weeks with IFN- $\alpha$  or with a HCV inhibitory compound. Analysis at the end of the treatment showed that neither viral proteins nor replicon RNA could be detected.

Cured cells (10IFN and Cl.60/cu) were transfected with mutated replicons and replication efficiency was determined by counting neomycin resistant clones (10IFN) or by TaqMan (10IFN and Cl.60/cu). As shown in Table 8, for all tested replicons the G418 cfu efficiency in 10IFN cells was at least 5 fold higher than in parental Huh-7 cells. This increase in G418 cfu efficiency was particularly relevant for a subset of mutants (m3, m5, m8, m9, m15).

**TABLE 8**

| Construct     | Consensus mutations |                |           | G418 cfu/10 <sup>3</sup> transfected cells |
|---------------|---------------------|----------------|-----------|--|
|               | NS3                 | NS5A           | EMCV IRES |  |
| pHCVNeo17.wt  |                     |                |           | 12 - 56                                    |
| pHCVNeo17.GAA |                     |                |           | 0  |
| pHCVNeo17.m0  |                     | S2204R         |           | 180 - 1000                                 |
| pHCVNeo17.m1  |                     | N2041T         |           | 8 - 13                                     |
| pHCVNeo17.m2  |                     | S2173F         |           | 2000                                       |
| pHCVNeo17.m3  |                     | S2197F         |           | 1600 - 3000                                |
| pHCVNeo17.m4  |                     | L2198S         |           | 190 - 650                                  |
| pHCVNeo17.m5  |                     | K@2039         |           | 1600 - 3000                                |
| pHCVNeo17.m6  | E1202G; A1347T      | S2173F         | extra A   | 600 - 2000                                 |
| pHCVNeo17.m7  |                     | N2041T; S2173F |           | 170 - 800                                  |
| pHCVNeo17.m8  |                     | N2041T; S2197F |           | > 4000                                     |
| pHCVNeo17.m9  |                     | N2041T; L2198S |           | 1400 - 3000                                |
| pHCVNeo17.m10 | E1202G              | K@2039         |           | >4000                                      |
| pHCVNeo17.m11 |                     | S2197F; A2199T |           | >4000                                      |



**TABLE 8**

| Construct     | Consensus mutations         |      |           | G418 cfu/10 <sup>5</sup><br>transfected cells |
|---------------|-----------------------------|------|-----------|---|
|               | NS3                         | NS5A | EMCV IRES |   |
| pHCVNeo17.m12 | N2041T; S2197F; A2199T      |      |           | >4000   |
| pHCVNeo17.m13 | N2041T; S2197F      extra A |      |           | >4000   |
| pHCVNeo17.m14 | S2197F; A2199T      extra A |      |           | >4000   |
| pHCVNeo17.m15 | A2199T                      |      |           | > 4000  |

10IFN cells (2x10<sup>6</sup>) were transfected with 10 µg of RNA transcribed from the indicated constructs. Approximately 2x10<sup>5</sup> cells were plated in a 10 cm tissue culture dish and cultured with 1 mg/ml G418 for 20 days.

Colonies surviving selection were stained with crystal violet and counted.

Strikingly, the best mutants yielded a number of G418 resistant clones ranging between 20 and 80% of the cell clones which grew in the absence of G418 (data not shown), thus indicating that the majority of 10IFN cells were competent for replication. This result was confirmed by TaqMan analysis (Table 9), in which the fold increase versus the parental Huh-7 cells was very high. The data indicates that replicons carrying adaptive mutations replicate vigorously in replicon enhanced cells such as 10IFN and Cl.60/cu.

**TABLE 9**

| Name         | 10IFN    |          | Cl.60/cu. |          |
|--------------|----------|----------|-----------|----------|
|              | RNA equ. | RNA equ. | RNA equ.  | RNA equ. |
|              | Day 4    | day 6    | day 4     | Day 6    |
| Wt           | 1 x      | 1 x      | 1 x       | 1 x      |
| hcvneo17.m0  | 46 x     | 12 x     | 78 x      | 512 x    |
| hcvneo17.m2  | 2 x      | 2 x      | 1 x       | 2 x      |
| hcvneo17.m3  | 68 x     | 49 x     | 19 x      | 392 x    |
| hcvneo17.m15 | 247 x    | 80 x     | 268 x     | 5518 x   |

Expression of viral proteins was determined in replicon enhanced cells using an ELISA assay designed to detect the NS3 protein in transfected cells plated in 96 wells microtiter plates (Cell-ELISA). As shown in Table 10, 24 hours post-transfection cells transfected with all tested replicons expressed low but detectable levels of the NS3 protein.

**TABLE 10**

| Name          | NS3 arbitrary units |       |           |      |
|---------------|---------------------|-------|-----------|------|
|               | 24 h p.t.           |       | 96 h p.t. |      |
|               | -                   | + IFN | -         | +IFN |
| Construct     |                     |       |           |      |
| Mock          | 1                   | 1     | 1         | 1    |
| pHCVNeo17.wt  | 3.7                 | 4.2   | 1.2       | 1.3  |
| pHCVNeo17.GAA | 3.1                 | 3.2   | 1.1       | 1    |
| pHCVNeo17.m0  | 3.4                 | 3.2   | 9.9       | 0.8  |
| pHCVNeo17.m3  | 5.7                 | 4.6   | 4.7       | 1.5  |
| pHCVNeo17.m8  | 6.6                 | 5.1   | 15.1      | 1.4  |
| pHCVNeo17.m10 | 8                   | 5.6   | 9.2       | 1.8  |
| pHCVNeo17.m11 | 8.4                 | 6.2   | 13.6      | 1.8  |

10IFN cells ( $2 \times 10^6$ ) were transfected with 10 µg of RNA transcribed from the indicated constructs. Cells were plated in 96 wells microtiter plates as indicated in Example 1.

- 5 Where indicated (+IFN), IFN- $\alpha$  (100 U/ml) was added to the culture medium 4 hours post-transfection. At the indicated times post-transfection, cells were fixed and analyzed by Cell-ELISA.

The early expression shown in Table 10 is likely due to translation of transfected RNA, since it was comparable in all replicons (including that carrying the GAA mutation) and was not affected by IFN- $\alpha$ . At 4 days post-transfection, NS3  
 10 expression persisted or increased in cells transfected with replicons carrying consensus mutations, but could not be detected anymore in cells transfected with wt and GAA replicons. In addition, NS3 expression was almost completely abolished when cells were cultured in the presence of IFN- $\alpha$ .

Taken together, these results indicated that the level of NS3 expression  
 15 reflected the replication rate. Indeed, NS3 expression level (Table 10) paralleled the RNA level measured by TaqMan (Table 9). The high replication proficiency of 10IFN cells was further confirmed by immunofluorescence experiments which showed that more than 50% of cells transfected with replicons m8 and m11 expressed high level of viral proteins, and that expression was almost completely abolished by  
 20 IFN- $\alpha$ .

#### Example 7: Replication of Full Length Constructs

This example illustrates the ability of a full length HCV genome containing adaptive mutations described herein to replicate in a replicon enhanced  
 25 host cell. The full length sequence of the HCV isolate Con-1 (EMBL-Genbank No. AJ238799) (plasmid pHCVRBFL.wt) and 2 derivatives containing either the N2041T

and S2173 F mutations (plasmid pHCVRBFL.m8) or the S2197F and A2199T mutations (plasmid pHCVRBFL.m11) were used as starting constructs.

- 5 RNAs transcribed from the starting constructs were transfected in 10IFN cells and their replication proficiency was assessed by Cell-ELISA, immunofluorescence and TaqMan. Both constructs containing consensus mutations (pHCVRBFL.m8 and pHCVRBFL.m11) replicated, while no sign of replication was observed with the wt. construct (data not shown).

#### Example 8: Replicons with Reporter Gene

- 10 This example illustrates an HCV replicon containing adaptive mutations and a reporter gene. A pHCVNeo17.wt derivative where the Neo coding region was substituted with that coding for human placental secretory alkaline phosphatase (pRBSEAP5.wt) and a derivative also containing the N2041T and S2173F mutations (plasmid pRBSEAP5.m8) were constructed. RNAs transcribed  
15 from these plasmids were transfected in 10IFN cells and their replication proficiency was assessed by measuring secretion of alkaline phosphatase. Analysis of the kinetics of secretion suggested that only plasmid pRBSEAP5.m8 was competent for replication (data not shown).

#### 20 Example 9: SEQ. ID. Nos. 1 and 2

SEQ. ID. NOs. 1 and 2 are provided as follows:

##### SEQ. ID. NO. 1

- 25 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKT  
SERSQPRGRRQPIPKARQPEGRAWAQP GYPWPLYGNEGLGWAGWLLSPRGS  
RPSWGPTDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGV  
RVLEDGVNYATGNLPGCSFSIFLLALLSCLTIPASAYEVRNVSGVYHVTNDCS  
NASIVYEAADMIMHTPGCVPCVRENSSRCWVALTPTLAARNASVPTTTIRR  
HVDLLVGAAALCSAMYVGDLGCVFLVAQLFTFSPRRHETVQDCNCSIYPGH  
30 VTGHRMAWDMMMNWSPTAALVVSQLLRIPQAVVDMVAGAHWGVLAGLA  
YYSMVGNWAKVLIVMLLFAGVDGGTYVTGGTMAKNTLGITSLFSPGSSQKIQ  
LVNTNGSWHINRTALNCNDSLNTGFLAALFYVHKFNSSGCPERMASCSPIDAF  
AQQWGPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVC GPVYCFT PSPVVVG  
TTDRFGVPTYSWGENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGGPP

CNIGGIGNKTLTCPTDCFRKHPEATYTKCGSGPWLTTPRCLVHYPYRLWHYPC  
TVNFTIFKVRMYVGGVEHRLEAACNWTRGERCNLEDRDRSELSPLLLSTTEW  
QVLPCSFTTLPALSTGLIHLHQNVDVQYLYGIGSAVVSFAIKWEYVLLLFLLL  
ADARVCACLWMMLLIAQAEAALENLVVNAASVAGAHGILSFLVFFCAAWY  
5 IKGRLVPGAAYALYGVWPLLLLLLALPPRAYAMDREMAASCGGAVFVGLILL  
TLSPHYKLFLARLIWWLQYFITRAEAHLQVWIPPLNVRGGRDAVILLTCAIHPE  
LIFTITKILLAILGPLMVLQAGITKVPYFVRAHGLIRACMLVRKVAGGHYVQM  
ALMKLAALTGTYYVDHLTPLRDWAHAGLRDLAVAVEPVVFSDMETKVITW  
GADTAACGDIILGLPVSARRGREIHLGPADSLEGQGWRLAPITAYSQQTRGL  
10 LGCIITSLTGRDRNQVEGEVQVSTATQSFLATCVNGVCWTVYHGAGSKTLA  
GPKGPIQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVR  
RRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFV  
PVESMETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSKGSTKVPAAYAAQG  
YKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYGFKA  
15 DGGCSGGAYDIICDECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSV  
TVPHPNIEEVALSSTGEIPFYGKAIPETIKGGRHLIFCHSKKKCEDELA AKLSGLG  
LNAVAYYRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSVIDCNTCVTQTV  
FSLDPTFTIETTTVPQDAVSRSQRRGRTGRGRMGIYRFVTPGERPSGMFDSV  
CECYDAGCAWYELTPAETSVRLRAYLNTPLPVCQDHLEFWESVFTGLTHID  
20 AHFLSQTKQAGDNFPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKPTLHG  
PTPLLYRLGAVQNEVTTHPITKYIMACMSADLEVVTSTWVLVGGVLAALAA  
YCLTTGSVVIVGRILSGKPAIPDREVL YREFDEMEECASHLPYIEQGMQLAEQ  
FKQKAIGLLQTATKQAEAAAPVVESKWR TLEAFWAKHMWNFISGIQYLAGLS  
TLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAG  
25 IAGAAVGSIGLGKVLVDILAGYGAGVAGALVAFKVMMSGEMPSTEDLVNLLPA  
ILSPGALVVGVC AAILRRHVGPGE GAVQWMNRLIAFASRGNHVSPTHYVPE  
SDAAARVTQILSSLTITQLLKRLHQWINE DCSTPCSGSWLRD VWDWICTVLTD  
FKTWLQSKLLPRLPGVPFFSCQRGYKGVWRGDGIMQTTCPGQAQITGHVKNG  
SMRIVGPRTCSNTWHGTFPINA YTTGPCTPSPAPNYSRALWRVAAEEYVEVT  
30 RVGDFHYVTGMTTDNVKCPQVPAPEFFTEVDGVRLHRYAPACKPLLREEV  
TFLVGLNQYLVGSQLPCEPEPDVA VLTSMLTDPSHITAETA KRRLARGSPPSL  
ASSASQLSAPSLKATCTTRHDS PDADLIEANLLWRQEMGGNITRVESENKV  
ILDSFEPLQAEEDEREVS VPAEILRRSRKFPRAMPIWARPDYNPPLLESWKDPD  
YVPPVVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALAE LATKTFGSSESA  
35 VDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPGDPDLSDGSWSTVSEE

ASDEVVCCSMSYTWGTALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSA  
 SLRQKKVTFDRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHS  
 ARSKFGYGA KDVRNLSSKAVNHRSVWKDLLED TETPIDTTIMAKNEVFCVQ  
 PEKGG RKPRLIVFPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQR  
 5 VEFLVNAWKAKKCPMGFAYDTRCFDSTVTENDIRVEESIQCCDLAPPEARQA  
 IRS LTERLYIGGPLTNSKGQNCGYRRCRASGVLTTS CGNTLT CYLKAAAACRA  
 AKLQDCTMLVCGDDL VICESAGTQEDEASLRAFTEAMTRY SAPPGDPPKPE  
 YDLELITSCSSNVSAHDASGKRVYYLTRDPTT PLARAAWETARHTPVNSWL  
 GNIIMYAPTLWARMILMTHFFSILLAQE QLEKALDCQIYGACYSIEPLDLPQIIQ  
 10 RLHGLSAFSLHSYSPGEINRVASCLRKLGVPLRVWRHRARSVRARLLSQGGR  
 AATCGKYLFNWAVRTKLKLTPIPAASQLDLSSWFVAGYSGGDIYHSLSRARP  
 RWF MWCLLLSVGVGIYLLPNR

SEQ. ID. NO. 2:

15 gccagccccgattggggcgacactccaccatagatcactcccctgtgaggaactactgtcttcacgcagaaagcgtcta  
 gccatggcgtagtagtgagtgctgcagcctccaggacccccctccgggagagccatagtggtctgcggaaccgggtg  
 agtacaccggaattgccaggacgaccgggtcctttcttgatcaaccgcctcaatgcctggagattgggcgtgccccgcg  
 agactgctagccgagtagtggtgcgaaaggccttggtactgcctgatagggtgctgcgagtgccccgggagggt  
 ctcgtagaccgtgcaccatgagcacgaatcctaaccctcaaaaaaccaaacgtaacaccaaccgcccccacagga  
 20 cgtcaagttccccggcggtggtcagatcgtcggtggagttacctgttgccgcgcagggggccccaggttggtggtgcgcgc  
 gactaggaagacttccgagcgggtcgcaacctcgtggaaggcgacaacctatcccaaggctcgcagcccagggtagg  
 gcctgggtcagccccgggtacccctgccccctatggcaatgagggttgggggtgggcaggatggctcctgtacccccgt  
 ggctctcgccatagttggggccccacggacccccggcgtaggctgcgcaatttgggtaagggtcatcgaacctcacgtgc  
 ggcttcgccgatctcatgggtacattccgctcgtcgccgccccctagggggcgtgccagggccctggcgcatggcgt  
 25 ccgggttctggaggacggcgtgaactatgaacagggaatctgccgggtgctcctttctatcttcttttggctttgtgtcct  
 gtttgaccatcccagcttccgcttatgaagtgcgcaacgtatccggagtgtagcatgtcacgaacgactgtccaacgcaag  
 cattgtgatgaggcagcggacatgatcatgataccccgggtgcgtgccctgcgttcgggagaactcctccccgtgc  
 tgggtagcgtcactcccacgctcgcggccaggaacgtagcgtccccactacgacgatacagccatgtcgattgtc  
 gttggggcggctgctctgtctccgtatgtacgtgggagatctctcggtatctgtttcctcgtcggcagctgttcaccttctc  
 30 gcctcgccggcacgagacgtacaggactgcaattgctcaatatccccggccacgtgacaggtcaccgtatggctggga  
 tatgatgatgaactgggtcacctacagcagccctagtggtatcgagttactccggatcccacaagctgtcgtggatatggtg  
 cggggggccattggggagtcctagcgggccttgctactatccatgggtggggaactgggctaagggttctgattgtatgcta  
 ctcttggccggttgacgggggaacctatgtacaggggggacgatggccaaaaacacctcgggattacgtccctcttll  
 caccgggtcatcccagaaaatccagcttgaataaccaacggcagctggcacatcaacaggactgccctgaactgcaat  
 35 gactccctcaactgggttcttgcgtcgtgttctacgtgcacaagttcaactcatctggatgccagagcgcagtgccag

ctgcagccccatcgacgcgttcgctcaggggtgggggcccacattacaatgagtcacacagctcggaccagaggccitta  
ttgttgccactacgcaccccgccgtgcgggtatcgtaaccgcgccgaggtgtgtgtccagtgtactgctcaccccaagc  
cctgtcgtggtggggacgaccgaccgggttcggcgtccctacgtacagttggggggagaatgagacggacgtgctgcttctt  
aacaacacgcggccgcccgaaggcaactggttgggtgtacatggatgaatagcactgggttcaccaagcgtgcggggg  
5 ccccccgtgtaacatcggggggatcggcaataaaaccttgacctgccccacggactgctccggaagcaccocgaggcca  
cttacaccaagtgtggttcggggccttgggtgacacccagatgcttgggtccactaccatacaggcttggcactaccctgc  
actgtcaactttaccatctcaaggttaggatgtacgtggggggagtgaggcacaggctcgaagccgatcaattggactc  
gaggagagcgttgaacctggaggacaggacagatcagagcttagcccgtgctgctgtctacaacggagtggcaggtta  
ttgccctgttcccttaccacccctaccggctctgtccactggtttagtccatctccatcagaacgtcgtggacgtacaatactgt  
10 acggtatagggtcggcggttgtctcctttgcaatcaaatgggagtatgtcctgttgccttcttcttcttggcggacgcgcgcgt  
ctgtgcctgcttggatgatgctgctgatagctcaagctgaggccgccctagagaacctgggtggtcctcaacgcggcatcc  
gtggccggggcgcatggcattctctccttctctgttcttctgtgctgcctggtacatcaaggcgaggctggtccctggggc  
ggcatatgccctctacggcgtatggccgctactcctgctcctgctggcgttaccaccacgagcatacgccatggaccggga  
gatggcagcatcgtgcggaggcgcggttttctaggtctgatactttagccttgcaccgcactataagctgttccctcgttag  
15 gctcatatggtggttacaatatlttaccagggccgaggcacacttgcaagtgtggatccccccctcaacgttcgggggg  
gccgcgatgccgtacccctcacgtgcgcgatccaccagagctaattttaccatcaccaaaatcttgcctccatactc  
ggtccactcatggtgctcaggctggtataacaaaagtgcgtacttctgctgcgcacacgggctattcgtcatgcatgct  
ggtgcggaagggttgcgggggtcattatgtccaaatggctctcatgaagtggccgactgacaggtacgtacgtttatgacc  
atctcaccocactgcgggactggggccacgcgggcctacgagaccttgcgggtggcagttgagcccgctgcttctctgat  
20 ggagaccaaggttatcacctggggggcagacaccgcggcggtgtggggacatcatcttgggcctgcccgctccgccgca  
gggggagggagatacatctgggaccggcagacagccttgaaggcgagggtggcgactcctcgcgcctattacggccta  
ctcccaacagacgcgaggcctacttggctgcatcatcactagccctcacaggccgggacaggaaccaggctgagggggag  
gtccaagtggctccaccgcaacacaatcttctggcgacctgcgtcaatggcgtgtgttggactgtctatcatggtgccgg  
ctcaaaagaccttgcgggcccgaaggggcccaatcaccgaatgtgaccaggacctcgtcggctgggaag  
25 cggcccccggggcgcttcttgacacatgcacctgcggcagctcggaccttllacttgggtcacgaggcatgccgatgtcat  
tccggtgcggcgggggcgacagcagggggagcctactctccccaggcccgctcctacttgaagggtcttctgggc  
ggtccactgctctgcccctcggggcacgctgtgggcacatcttctgggctgccgtgtgcacccgaggggttgcgaaggcgggtg  
gacttltacccgtcagctctatggaaaccactatcggtlccccggcttctacggacaactcgtcccctccggccgtaccgc  
agacattccagggtggcccatctacacgcccctactggtagcggcaagagcactaagggtccgggtcgtatgcagcccaa  
30 ggggtataagggtgcttgcctgaacccgtcgtccgccaccctaggtttcggggcglatatgtctaaggcacatggtatcga  
ccctaacatcagaaccggggtaaggaccatcaccacgggtgcccccatcacgtactccacctatggcaagtttcttgcggac  
gggtggttgccttggggggcgccctatgacatcataatatgtgatgagtgccactcaactgactcgaccactatcctgggcatcgg  
cacagtcttgaccaagcggagacggctggagcgcgactcgtcgtcgtccaccgctacgcctccgggatcgggtcacc  
gtgccacatccaaacatcgaggagggtggctctgtccagcactggagaaatccccctttatggcaagccatccccatcgaga  
35 ccatcaagggggggaggcacctcatlttctgccattccaagaagaaatgtgatgagctcggcggaagctgtccggcctcg

gactcaatgctgtagcatattaccggggccttgatgtatccgtcataccaactagcggagacgtcattgtcgtagcaacggac  
 gctctaataacgggctttaccggcgatttcgactcagtgatcgactgcaatacatgtgtcaccagacagtcgacttcagcct  
 ggacccgaccttcaccattgagacgacgaccgtgccacaagacgcgggtgcacgctcgcagcggcgaggcaggactggg  
 agggggcaggatgggcatttacaggtttgtgactccaggagaacggccctcgggcatgttcgattctcggttctgtgcgagt  
 5 gctatgacgcgggctgtgcttggtacgagctcacgcccgcgagacctcagttaggttgcgggcttacctaaacacaccag  
 ggttgcccgctgccaggaccatctggagttctgggagagcgtctttacaggccacccacatagacgccatttctgtcc  
 cagactaagcaggcaggagacaactccccctacctggtagcataccaggctacgggtgtgccaggggctcaggctccacc  
 tccatcgtgggaccaaagtgggaagtgtctacacggctaaagcctacgtgcacgggccaacggccctgctgtataggctg  
 ggagccgttcaaaacgaggttactaccacacacccataaccaatacatcatggcatgcatgtcgggtgacctggaggtc  
 10 gtcacgagcacctgggtgtgtagggcggagtcctagcagctctggccgcgtattgcctgacaacaggcagcgtggtcatt  
 gtgggcaggatcatctgtccggaaagccggccatcattcccgacagggaagtcctttaccgggagttcgaatgagatggaa  
 gagtgcgcctcacacctcccttacaatgaacagggaatgcagctcgcgaacaattcaaacagaaggcaatcgggttctg  
 caaacagccaccaagcaagcggaggctgtgtctccgtggtggaatccaagtggcggaccctcgaagccttctgggcga  
 agcatatgtggaatttcacagcgggatacaatatttagcaggcttgcctcctggcaaccccgcatagcatcactga  
 15 tggcaticacagcctctatcaccagcccgtcaccaccaacataccctcctgtttaacatcctggggggatgggtggccgc  
 ccaacttgctcctcccagcgtgcttctgtttcgtagggcggcggcatcgtggagcggctgttggcagcataggccttggg  
 aaggtgcttggatattttggcaggttatggagcaggggtggcaggcgcgctcgtggcctttaaggtcatgagcggcgag  
 atgccctccaccgaggacctggttaacctactccctgtatcctctccctggcggcctagtcgtcggggtcgtgtgcgcagc  
 gatactgcgtcggcacgtgggcccaggggagggggctgtgcagtggaatgaaccggctgatagcgttcgcttcgcgggga  
 20 accacgtcctcccccacgcactatgtgctgagagcgacgctgcagcacgtgtcactcagatcctctctagcttaccatcact  
 cagctgctgaagaggcttcaccagtggaatcaacgaggactgtccacgccatgctccggctcgtggctaagagatgtttgg  
 gattggatagcacggtgttgactgattcaagacctggctccagtcgaagctcctgccgcgattgccgggagtccttcttc  
 tcatgtcaacgtgggtacaaggagctgtggcggggcgacggcatcatgcaaacacctgcccattgtggagcacagatcac  
 cggacatgtgaaaaacgggtccatgaggatcgtggggcctaggacctgtagtaacacgtggcatggaacattccccattaac  
 25 gcgtacaccacggggccctgcacgccctccccggcgccaaattattctaggcgctgtggcgggtggctgctgaggagta  
 cgtggagggttacgcgggtgggggatttccactacgtgacgggcatgaccactgacaacgtaaaagtcccgtgtcaggtcc  
 ggccccgaattcttcacagaagtggatgggtgcgggtgcacagggtacgtccagcgtgcaaacccctctacgggagg  
 aggtcacattcctggctgggctcaatcaatacctggttgggtcacagctcccatgcgagcccgaaccggacgtagcagtgct  
 cacttccatgctcaccgacccctcccacattacggcgagacggctaagcgtaggctggccaggggatctccccctcctt  
 30 ggccagctcatcagctagccagctgtctgcgccttcttgaaggcaacatgcactaccgctcatgactccccggacgctgac  
 ctcatcaggccaacctcctgtggcggcaggagatggcggggaacatcacccgcgtggagtcagaaaataaggtagtaaat  
 ttggactcttctcagccgctccaagcggaggaggatgagaggggaagtatccgttccggcgagatcctgcggagggtcca  
 ggaaattccctcgagcgtatgccatattgggcacggcggattacaacctccactgttagagtcctggaaggacccggact  
 acgtccctccagtggtacacgggtgtccattgccgctgccaaggccctccgataccacctccacggagggaaggagacg  
 35 gttgtcctgtcagaatctaccgtgtcttctgccttggcgaggctcggcacaagaccttcggcagctccgaatcgtcggcgt

cgacagcggcagcggcaacggcctctcctgaccagccctccgacgacggcgacgcgggatccgacgttgagtcgtactcc  
 tccatgcccccccttgagggggagccgggggatcccgatctcagcgacgggtcttggtctaccgtaagcgaggaggctag  
 tgaggacgtcgtctgctgctcgaatgtctacacatggacagcgccctgatcacgccatgcgctgcggaggaaaccaagct  
 gccccatcaatgcactgagcaactctttgctccgicaccacaacttggctatgctacaacatctcgacgcgaagcctgcggc  
 5 agaagaaggcacccttgacagactgcaggctctggacgaccactaccgggacgtgctcaaggagatgaaggcgaaggc  
 gtccacagttaaggctaaacttctatccgtggaggaagcctgtaagctgacgccccacattcggccagatctaatttggct  
 atggggcaaaggacgtccggaacatccagcaaggccgtaaccacatccgctccgttggaaggacttctggaagac  
 actgagacaccaattgacaccacatcatggcaaaaaatgaggtttctgcgtccaaccagagaaggggggccgcaagcc  
 agctcgccttatcgtattcccagatttgggggtcgtgtgctgcgagaaaatggccctttacgatgtggtctccacctccctcag  
 10 gccgtgatgggctcttcatacggattccaatactctcctggacagcgggtcgagttcctggtgaatgcctggaaagcgaaga  
 aatgccctatgggcttcgcatatgacaccgcgttttgactcaacggctactgagaatgacatccgtgttgaggagtcaatct  
 accaatgttgtgacttggccccgaagccagacaggccataaggtcgctcacagagcgggtttacatcgggggccccctga  
 ctaattctaagggcagaactgcggctatcgccggtgccgcgcgagcgggtgactgacgaccagctgcggtaataccctca  
 catgttacttgaaggccgctcgggcctgtcgagctcggaagctccaggactgcacgatgctcgtatcgggagacgaccttgt  
 15 cgttatctgtgaaagcgcggggacccaagaggacgagcgagcctacggccttcacggaggctatgactagatactctg  
 cccccctggggacccgccccaaaccagaatacgacttggagttgataacatcatgctcctccaatgtgicagtcgcgcacg  
 atgcatctggcaaaagggtgactatctacccgtgacccaccaccccccttgcgcgggctgcgtgggagacagctagac  
 acactccagtcattcctggctaggcaacatcatcatglatgcgccaccttggggcaaggatgacctgatgactcatttctt  
 ctccatccttctagctcaggaacaacttgaaaaagccctagattgctcagatctacggggcctgttactccattgagccacttga  
 20 cctacctcagatcatlcaacgactccatggccttagcgcatcttactccatagttactctccaggtgagatcaatagggtggct  
 tcatgcctcaggaaacttgggtaccgcccttgcgagcttgagacatcgggccagaaggtccgcgctaggtactgtcc  
 cagggggggagggtgcccacttgggcaagtaccttcaactgggcagtaaggaccaagctcaactcactccaatccccg  
 gctgcgtcccagtlggatltatccagctgggtcgttgcgtgttacagcgggggagacatatcacagcctgtctcgtgccgga  
 ccccgctggttcatgtggtgcctactcctacttctgtagggttaggcacatctactctcccaaccgatgaacggggagctaa  
 25 acactccaggccaataggccatcctgttttttccctt  
 ttccttttttcttcttgggtggctccatcttagccctagtcacggctagctgtgaaaggctccgtgagccgcttgactgcagagagt  
 gctgatactggcctctctgcagatcaagt

Other embodiments are within the following claims. While several  
 30 embodiments have been shown and described, various modifications may be made  
 without departing from the spirit and scope of the present invention.



## WHAT IS CLAIMED IS:

1. A nucleic acid molecule comprising a region selected from the group consisting of:
  - 5 a) an altered HCV NS3 encoding region coding for one or more NS3 mutations, wherein at least one of said NS3 mutations, identified by reference to the amino acid sequence numbering of SEQ. ID. NO. 1, is selected from the group consisting of:
    - amino acid 1095 being Ala,
    - 10 amino acid 1202 being Gly, and
    - amino acid 1347 being Thr;
  - b) an altered HCV NS5A encoding region coding for one or more NS5A mutations, wherein at least one of said NS5A mutations, identified by reference to the amino acid sequence numbering of SEQ. ID. NO. 1, is selected from the group  
15 consisting of:
    - amino acid 2041 being Thr,
    - a Lys insertion between residue 2039 and 2040.
    - amino acid 2173 being Phe,
    - amino acid 2197 being Phe,
    - 20 amino acid 2198 being Ser,
    - amino acid 2199 being Thr, and
    - amino acid 2204 being Arg; and
  - c) an altered encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region containing one or more EMCV IRES mutations,  
25 wherein at least one of said EMCV IRES mutations, identified by reference to the nucleotide number of SEQ. ID. NO. 3, is an insertion at nucleotide 1736 of adenine.
2. The nucleic acid molecule of claim 1, wherein said nucleic acid molecule comprises said NS5A encoding region.  
30
3. The nucleic acid molecule of claim 2, wherein at least two of said NS5A adaptive mutations are present.

4. The nucleic acid molecule of claim 2, further comprising a region encoding for a HCV NS3 region, wherein said NS3 region may be the same or different than said altered NS3 region.

5 5. The nucleic acid molecule of claim 4, wherein said nucleic acid molecule is an HCV replicon comprising a HCV 5' UTR-PC region, said NS3 encoding region, an HCV NS4A encoding region, an HCV NS4B encoding region, said NS5A encoding region, an HCV NS5B encoding region, and a HCV 3' UTR.

10 6. The nucleic acid molecule of claim 5, wherein said HCV replicon further comprises a sequence encoding for a reporter protein.

7. The nucleic acid molecule of claim 5, wherein said HCV replicon further comprises a sequence encoding for a selection protein.

15 8. The nucleic acid molecule of claim 5, wherein said HCV replicon further comprises a HCV core encoding region, a HCV E1 encoding region, a HCV E2 encoding region, a HCV p7 encoding region, and a HCV NS2 encoding region.

20 9. A nucleic acid molecule comprising a region selected from the group consisting of:

a) an altered HCV NS3 encoding region containing one or more NS3 mutations, wherein at least one of said NS3 mutations, identified by reference to  
25 the nucleotide numbering of SEQ. ID. NO. 2, is selected from the group consisting of: nucleotide 3625 being cytosine, nucleotide 3946 being guanine, nucleotide 4380 being adenine,

b) an altered HCV NS5A encoding region containing one or more  
30 NS5A mutations, wherein at least one of said NS5A mutations, identified by reference to the nucleotide numbering of SEQ. ID. NO. 2, is selected from the group consisting of:

an insertion of 3 adenine residues between nucleotide 6458 and 6459,  
nucleotide 6463 being cytosine,  
35 nucleotide 6859 being thymine or uracil,

nucleotide 6931 being thymine or uracil,  
nucleotide 6934 being cytosine,  
nucleotide 6936 being adenine, and  
nucleotide 6953 being adenine or guanine; and

- 5                   c)     an altered encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region containing one or more EMCV IRES mutations, wherein at least one of said EMCV IRES mutations, identified by reference to the nucleotide number of SEQ. ID. NO. 3, is an insertion at nucleotide 1736 of adenine.

- 10                   10.     The nucleic acid molecule of claim 9, wherein said molecule comprises said altered NS5A encoding region, and the nucleotide sequence of said altered NS5A region is provided for by bases 6258-7598 of SEQ. ID. NO. 2, or the RNA version thereof, modified with one or more of said NS5A modifications selected from the group consisting of:

- 15     an insertion of 3 adenine residues between nucleotide 6458 and 6459,  
nucleotide 6463 being cytosine,  
nucleotide 6859 being thymine or uracil,  
nucleotide 6931 being thymine or uracil,  
nucleotide 6934 being cytosine,  
20     nucleotide 6936 being adenine, and  
nucleotide 6953 being adenine or guanine.

11.     The nucleic acid molecule of claim 10, wherein said molecule is an HCV replicon comprising a HCV 5' UTR-PC region, a modified HCV NS3-NS5B region, and a HCV 3' UTR, wherein said modified NS3-NS5B region  
25     comprises said altered NS5A region.

12.     The nucleic acid molecule of claim 11, wherein said 5' UTR-PC region is the RNA version of bases 1-377 of SEQ. ID. NO. 2 and said 3' UTR is  
30     the RNA version of bases 9374-9605 of SEQ. ID. NO. 2.

13.     The nucleic acid molecule of claim 10, wherein said molecule is an HCV replicon comprising a HCV 5' UTR-PC region, a modified HCV NS3-NS5B region, and a HCV 3' UTR, wherein  
35     said 5' UTR-PC region is the RNA version of bases 1-377 of SEQ. ID. NO. 2;

said 3' UTR is the RNA version of bases 9374-9605 of SEQ. ID. NO. 2; and  
 said modified NS3-NS5B region consists of the RNA version of bases 3420-9371 of  
 SEQ. ID. NO. 2 modified with one or more modifications selected from the group  
 consisting of:

- 5 nucleotide 4380 being adenine,  
 nucleotide 3625 being cytosine,  
 nucleotide 3946 being guanine,  
 an insertion of 3 adenine residues between nucleotide 6458 and nucleotide 6459,  
 nucleotide 6463 being cytosine,
  - 10 nucleotide 6859 being uracil,  
 nucleotide 6931 being uracil,  
 nucleotide 6934 being cytosine,  
 nucleotide 6936 being adenine, and  
 nucleotide 6953 being adenine or guanine.
  - 15
14. The nucleic acid molecule of claim 13, wherein said replicon is  
 a genomic replicon that further comprises the RNA version of nucleotides 378-3419  
 of SEQ. ID. NO. 2.
- 20 15. A nucleic acid molecule comprising the nucleic acid base  
 sequence of bases 1-7989 of SEQ. ID. NO. 3, or the RNA version thereof, consisting  
 of one or more different modifications selected from the group consisting of:  
 a) nucleotides 5335-5337 modified to code for arginine;  
 b) nucleotides 5242-5244 modified to code for phenylalanine;
  - 25 c) nucleotides 5314-5316 modified to code for phenylalanine;  
 d) nucleotides 5317-5319 modified to code for serine;  
 e) nucleotides coding for lysine inserted after nucleotide 4843;  
 f) nucleotides 2329-2331 modified to code for glycine, nucleotides 2764-2766  
 modified to code for threonine, nucleotides 5242-5244 modified to code for  
 30 phenylalanine, and an extra adenosine inserted after nucleotide 1736;  
 g) nucleotides 4846-4848 modified to code for threonine, and nucleotides 5242-5244  
 modified to modified to code for phenylalanine;  
 h) nucleotides 4846-4848 modified to code for threonine, and nucleotides 5314-5316  
 modified to code for phenylalanine;

- i) nucleotides 4846-4848 modified to code for threonine, and nucleotides 5317-5319 modified to code for serine;
- j) nucleotides 2329-2331 modified to code for glycine, and nucleotides coding for lysine inserted after nucleotides 4843;
- 5 k) nucleotides 5314-5316 modified to code for phenylalanine and nucleotides 5320-5322 modified to code for threonine;
- l) nucleotides 4846-4848 modified to code for threonine, nucleotides 5314-5316 modified to code for phenylalanine, and nucleotides 5320-5322 modified to code for threonine;
- 10 m) nucleotides 4846-4848 modified to code for threonine, nucleotides 5314-5316 modified to code for phenylalanine, and an extra adenosine inserted after nucleotide 1736; and
- n) nucleotides 5314-5316 modified to code for phenylalanine, nucleotides 5320-5322 modified to code for threonine, and an extra adenosine inserted after nucleotide 1736;
- 15 and
- o) nucleotides 5320-5322 modified to code for threonine.

16. The nucleic acid of claim 15, wherein said one or more different modifications is selected from the group consisting of:

- 20 a) C5337A;
- b) C5243T or U;
- c) C5315T or U;
- d) T or U5318C;
- e) AAA inserted after 4843;
- 25 f) A2330G, G2764A, C5243T or U, and adenosine inserted 1736;
- g) A4847C and C5243T or U;
- h) A4847C and C5315T or U;
- i) A4847C and T or U5318C;
- j) A2330G and AAA inserted after 4843;
- 30 k) C5315T or U and G5320A;
- l) A4847C, C5315T or U, and G5320A;
- m) A4847C, C5315T or U, and adenosine inserted 1736;
- n) C5315T or U, G5320A and adenosine inserted 1736; and
- o) G5320A.

17. The nucleic acid of claim 16, wherein said nucleic acid is RNA and comprises said nucleic acid base sequence.

5 18. The nucleic acid of claim 17, wherein said nucleic acid is RNA and consists of said nucleic acid base sequence.

19. An expression vector comprising a nucleotide sequence coding for the nucleic acid molecule of any one of claims 1-18, wherein said nucleotide  
10 sequence is transcriptionally coupled to an exogenous promoter.

20. A recombinant cell human hepatoma cell, wherein said cell comprises the nucleic acid of any one of claims 5-8 and 11-18.

15 21. The recombinant cell of claim 20, wherein said hepatoma cell is an Huh-7 cell.

22. The recombinant cell of claim 20, wherein said cell is derived from a Huh-7 cell.  
20

23. A recombinant cell made by a process comprising the step of introducing into a human hepatoma cell the nucleic acid of any one of claims 5-8 and 11-18.

25 24. A method of making an HCV replicon enhanced cell comprising the steps of:  
a) introducing and maintaining a HCV replicon in a cell; and  
b) curing said cell of said HCV replicon to produce said replicon enhanced cell.  
30

25. The method of claim 24, wherein said cell is a human hepatoma cell.

26. The method of claim 24, wherein said cell is a Huh-7 cell or is  
35 derived from a Huh-7 cell.

27. The method of claim 26, further comprising the step of confirming the ability of said replicon enhanced cell to maintain an HCV replicon.

5                   28     A method of making an HCV replicon enhanced cell containing a functional HCV replicon comprising the steps of:

- a)     introducing and maintaining a first HCV replicon in a cell;
- b)     curing said cell of said first replicon to produce a cured cell;

and

10                   c)     introducing and maintaining a second HCV replicon into said cured cell, wherein said second HCV replicon may be the same or different than said first HCV replicon.

15                   29     The method of claim 28, wherein said cell is a human hepatoma cell.

30.     The method of claim 29, wherein said human hepatoma cell is a Huh-7 cell.

20                   31.     The method of claim 30, wherein said human hepatoma cell is derived from a Huh-7 cell.

32.     An HCV replicon enhanced cell made by the method of any one of claims 24-27.

25

33.     An HCV replicon enhanced cell containing a HCV replicon made by the method of any one of claims 28-31.

30                   34.     A method of measuring the ability of a compound to affect HCV activity comprising the steps of:

- a)     providing said compound to the HCV replicon enhanced cell of claim 33; and
- b)     measuring the ability of said compound to effect one or more replicon activities as a measure of the effect on HCV activity.

35

35. The method of claim 34, wherein said compound is a ribozyme.
36. The method of claim 34, wherein said compound is an  
antisense nucleic acid.
- 5 37. The method of claim 34, wherein compound is an organic  
compound.
- 10 38. The method of claim 34, wherein said step (b) measures HCV  
protein production.
39. The method of claim 33, wherein said step (b) measures  
production of RNA transcripts.



1 GCCAGCCCC GATTGGGGC GACACTCCAC CATAGATCAC TCCCCTGTGA  
51 GGAACTACTG TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG  
101 TGTCGTGCAG CCTCCAGGAC CCCCCCTCCC GGGAGAGCCA TAGTGGTCTG  
151 CGGAACCGGT GAGTACACCG GAATTGCCAG GACGACCGGG TCCTTTCTTG  
201 GATCAACCCG CTCAATGCCT GGAGATTGG GCGTGCCCC GCGAGACTGC  
251 TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG  
301 GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGAC CATGAGCACG  
351 AATCCTAAAC CTCAAAGAAA AACCAAAGGG CGCGCCATGA TTGAACAAGA  
401 TGGATTGCAC GCAGGTTCTC CGGCCGCTTG GGTGGAGAGG CTATTGCGCT  
451 ATGACTGGGC ACAACAGACA ATCGGCTGCT CTGATGCCGC CGTGTTCCGG  
501 CTGTACGCGC AGGGGCGCCC GGTCTTTTT GTCAAGACCG ACCTGTCCGG  
551 TGCCCTGAAT GAACTGCAGG ACGAGGCAGC GCGGCTATCG TGGCTGGCCA  
601 CGACGGGCGT TCCTTGCGCA GCTGTGCTCG ACGTTGTCAC TGAAGCGGGA  
651 AGGGA CTGGC TGCTATTGGG CGAAGTGCCG GGGCAGGATC TCCTGTGTCATC  
701 TCACCTTGCT CCTGCCGAGA AAGTATCCAT CATGGCTGAT GCAATGCGGC  
751 GGCTGCATAC GCTTGATCCG GCTACCTGCC CATTCGACCA CCAAGCGAAA  
801 CATCGCATCG AGCGAGCACG TACTCGGATG GAAGCCGGTC TTGTCGATCA  
851 GGATGATCTG GACGAAGAGC ATCAGGGGCT CGCGCCAGCC GAACTGTTGCG  
901 CCAGGCTCAA GGC GCGCATG CCCGACGGCG AGGATCTCGT CGTGACCCAT  
951 GGCGATGCCT GCTTGCCGAA TATCATGGTG GAAAATGGCC GCTTTTCTGG  
1001 ATTCATCGAC TGTGGCCGGC TGGGTGTGGC GGACCGCTAT CAGGACATAG  
1051 CGTTGGCTAC CCGTGATATT GCTGAAGAGC TTGGCGGCGA ATGGGCTGAC  
1101 CGCTTCCTCG TGCTTTACGG TATCGCCGCT CCCGATTGCG AGCGCATCGC  
1151 CTTCTATCGC CTTCTTGACG AGTTCTTCTG AGTTTAAACA GACCACAACG  
1201 GTTTCCCTCT AGCGGGATCA ATTCCGCCCC TCTCCCTCCC CCCCCCTAA  
1251 CGTTACTGGC CGAAGCCGCT TGGAAATAAG CCGGTGTGCG TTTGTCTATA  
1301 TGTTATTTTC CACCATATTG CCGTCTTTTG GCAATGTGAG GGCCCGGAAA  
1351 CCTGGCCCTG TCTTCTTGAC GAGCATTCCT AGGGGTCTTT CCCCTCTCGC  
1401 CAAAGGAATG CAAGGTCTGT TGAATGTCGT GAAGGAAGCA GTTCCTCTGG  
1451 AAGCTTCTTG AAGACAAACA ACGTCTGTAG CGACCCTTTG CAGGCAGCGG  
1501 AACCCCCAC CTGGCGACAG GTGCCTCTGC GGCCAAAAGC CACGTGTATA

FIG. 1A

1551 AGATACACCT GCAAAGGCGG CACAACCCCA GTGCCACGTT GTGAGTTGGA  
1601 TAGTTGTGGA AAGAGTCAAA TGGCTCTCCT CAAGCGTATT CAACAAGGGG  
1651 CTGAAGGATG CCCAGAAGGT ACCCCATTGT ATGGGATCTG ATCTGGGGCC  
1701 TCGGTGCACA TGCTTTACAT GTGTTTAGTC GAGGTTAAAA AACGTCTAGG  
1751 CCCCCGAAC CACGGGGACG TGGTTTTCCT TTGAAAAACA CGATAATACC  
1801 ATGGCGCCTA TTACGGCCTA CTCCCAACAG ACGCGAGGCC TACTTGGCTG  
1851 CATCATCACT AGCCTCACAG GCCGGGACAG GAACCAGGTC GAGGGGGAGG  
1901 TCCAAGTGGT CTCCACCGCA ACACAATCTT TCCTGGCGAC CTGCGTCAAT  
1951 GGCGTGTGTT GGACTGTCTA TCATGGTGCC GGCTCAAAGA CCCTTGCCGG  
2001 CCCAAAGGGC CCAATCACCC AAATGTACAC CAATGTGGAC CAGGACCTCG  
2051 TCGGCTGGCA AGCGCCCCC GGGGCGCGTT CCTTGACACC ATGCACCTGC  
2101 GGCAGCTCGG ACCTTTACTT GGTCACGAGG CATGCCGATG TCATTCCGGT  
2151 GCGCCGGCGG GCGGACAGCA GGGGGAGCCT ACTCTCCCC AGGCCCGTCT  
2201 CCTACTTGAA GGGCTCTTCG GGCGGTCCAC TGCTCTGCCC CTCGGGGCAC  
2251 GCTGTGGGCA TCTTTCGGGC TGCCGTGTGC ACCCGAGGGG TTGCGAAGGC  
2301 GGTGGACTTT GTACCCGTCG AGTCTATGGA AACCCTATG CGGTCCCCGG  
2351 TCTTCACGGA CAACTCGTCC CCTCCGGCCG TACCGCAGAC ATTCCAGGTG  
2401 GCCCATCTAC ACGCCCCTAC TGGTAGCGGC AAGAGCACTA AGGTGCCGGC  
2451 TGCGTATGCA GCCCAAGGGT ATAAGGTGCT TGTCTGAAC CCGTCCGTCG  
2501 CCGCCACCCT AGGTTTCGGG GCGTATATGT CTAAGGCACA TGGTATCGAC  
2551 CCTAACATCA GAACCGGGGT AAGGACCATC ACCACGGGTG CCCCATCAC  
2601 GTACTCCACC TATGGCAAGT TTCTTGCCGA CGGTGGTTGC TCTGGGGGCG  
2651 CCTATGACAT CATAATATGT GATGAGTGCC ACTCAACTGA CTCGACCACT  
2701 ATCTGGGCA TCGGCACAGT CCTGGACCAA GCGGAGACGG CTGGAGCGCG  
2751 ACTCGTCGTG CTCGCCACCG CTACGCCTCC GGGATCGGTC ACCGTGCCAC  
2801 ATCCAAACAT CGAGGAGGTG GCTCTGTCCA GCACTGGAGA AATCCCCTTT  
2851 TATGGCAAAG CCATCCCCAT CGAGACCATC AAGGGGGGGA GGCACCTCAT  
2901 TTTCTGCCAT TCCAAGAAGA AATGTGATGA GCTCGCCGCG AAGCTGTCCG  
2951 GCCTCGGACT CAATGCTGTA GCATATTACC GGGGCCCTGA TGTATCCGTC  
3001 ATACCAACTA GCGGAGACGT CATTGTCGTA GCAACGGACG CTCTAATGAC  
3051 GGGCTTTACC GCGATTTCG ACTCAGTGAT CGACTGCAAT ACATGTGTCA

FIG. 1B

3101 CCCAGACAGT CGACTTCAGC CTGGACCCGA CCTTCACCAT TGAGACGACG  
3151 ACCGTGCCAC AAGACGCGGT GTCACGCTCG CAGCGGCGAG GCAGGACTGG  
3201 TAGGGGCAGG ATGGGCATTT ACAGGTTTGT GACTCCAGGA GAACGGCCCT  
3251 CGGGCATGTT CGATTCTCTG GTTCTGTGCG AGTGCTATGA CGCGGGCTGT  
3301 GCTTGGTACG AGCTCACGCC CGCCGAGACC TCAGTTAGGT TGCGGGCTTA  
3351 CCTAAACACA CCAGGGTTGC CCGTCTGCCA GGACCATCTG GAGTTCCTGG  
3401 AGAGCGTCTT TACAGGCCTC ACCCACATAG ACGCCCATTT CTTGTCCCAG  
3451 ACTAAGCAGG CAGGAGACAA CTTCCCCTAC CTGGTAGCAT ACCAGGCTAC  
3501 GGTGTGCGCC AGGGCTCAGG CTCCACCTCC ATCGTGGGAC CAAATGTGGA  
3551 AGTGTCTCAT ACGGCTAAAG CCTACGCTGC ACGGGCCAAC GCCCCTGCTG  
3601 TATAGGCTGG GAGCCGTTCA AAACGAGGTT ACTACCACAC ACCCCATAAC  
3651 CAAATACATC ATGGCATGCA TGTGCGCTGA CCTGGAGGTC GTCACGAGCA  
3701 CCTGGGTGCT GGTAGGCGGA GTCCTAGCAG CTCTGGCCGC GTATTGCCTG  
3751 ACAACAGGCA GCGTGGTCAT TGTGGGCAGG ATCATCTTGT CCGGAAAGCC  
3801 GGCCATCATT CCCGACAGGG AAGTCCTTTA CCGGGAGTTC GATGAGATGG  
3851 AAGAGTGCGC CTCACACCTC CTTACATCG AACAGGGAAT GCAGCTCGCC  
3901 GAACAATTCA AACAGAAGGC AATCGGGTTG CTGCAAACAG CCACCAAGCA  
3951 AGCGGAGGCT GCTGCTCCCG TGGTGAATC CAAGTGGCGG ACCCTCGAAG  
4001 CCTTCTGGGC GAAGCATATG TGGAATTTCA TCAGCGGGAT ACAATATTTA  
4051 GCAGGCTTGT CCACTCTGCC TGGCAACCCC GCGATAGCAT CACTGATGGC  
4101 ATTCACAGCC TCTATCACCA GCCCGCTCAC CACCCAACAT ACCCTCCTGT  
4151 TTAACATCCT GGGGGGATGG GTGGCCGCC AACTTGCTCC TCCCAGCGCT  
4201 GCTTCTGCTT TCGTAGGCGC CGGCATCGCT GGAGCGGCTG TTGGCAGCAT  
4251 AGGCCTTGGG AAGGTGCTTG TGGATATTTT GGCAGGTTAT GGAGCAGGGG  
4301 TGGCAGGCGC GCTCGTGGCC TTTAAGGTCA TGAGCGGCGA GATGCCCTCC  
4351 ACCGAGGACC TGGTTAACCT ACTCCCTGCT ATCCTCTCCC CTGGCGCCCT  
4401 AGTCGTGCGG GTCGTGTGCG CAGCGATACT GCGTCGGCAC GTGGGCCAG  
4451 GGGAGGGGGC TGTGCAGTGG ATGAACCGGC TGATAGCGTT CGCTTCGCGG  
4501 GGTAACCACG TCTCCCCAC GCACTATGTG CCTGAGAGCG ACGCTGCAGC  
4551 ACGTGTCACT CAGATCCTCT CTAGTCTTAC CATCACTCAG CTGCTGAAGA  
4601 GGCTTCACCA GTGGATCAAC GAGGACTGCT CCACGCCATG CTCCGGCTCG

FIG. 1C

4651 TGGCTAAGAG ATGTTTGGGA TTGGATATGC ACGGTGTTGA CTGATTTCAA  
4701 GACCTGGCTC CAGTCCAAGC TCCTGCCGCG ATTGCCGGA GTCCCCTTCT  
4751 TCTCATGTCA ACGTGGGTAC AAGGGAGTCT GCGGGGCGA CGGCATCATG  
4801 CAAACCACCT GCCCATGTGG AGCACAGATC ACCGGACATG TGAAAAACGG  
4851 TTCCATGAGG ATCGTGGGGC CTAGGACCTG TAGTAACACG TGGCATGGAA  
4901 CATTCCCCAT TAACGCGTAC ACCACGGGCC CCTGCACGCC CTCCCCGGCG  
4951 CCAAATTATT CTAGGGCGCT GTGGCGGGTG GCTGCTGAGG AGTACGTGGA  
5001 GGTACGCGG GTGGGGGATT TCCACTACGT GACGGGCATG ACCACTGACA  
5051 ACGTAAAGTG CCCGTGTCAG GTTCCGGCCC CCGAATTCTT CACAGAAGTG  
5101 GATGGGGTGC GGTGTCACAG GTACGCTCCA GCGTGCAAAC CCCTCCTACG  
5151 GGAGGAGGTC ACATTCCTGG TCGGGCTCAA TCAATACCTG GTTGGGTCAC  
5201 AGCTCCCATG CGAGCCCGAA CCGGACGTAG CAGTGCTCAC TTCCATGCTC  
5251 ACCGACCCCT CACACATTAC GGCGGAGACG GCTAAGCGTA GGCTGGCCAG  
5301 GGGATCTCCC CCCTCCTTGG CCAGCTCATC AGCTAGCCAG CTGTCTGCGC  
5351 CTTCTTGAA GGCAACATGC ACTACCCGTC ATGACTCCCC GGACGCTGAC  
5401 CTCATCGAGG CCAACCTCCT GTGGCGGCAG GAGATGGGCG GGAACATCAC  
5451 CCGCGTGGAG TCAGAAAATA AGGTAGTAAT TTTGGACTCT TTCGAGCCGC  
5501 TCCAAGCGGA GGAGGATGAG AGGGAAGTAT CCGTTCCGGC GGAGATCCTG  
5551 CGGAGGTCCA GGAAATTCCC TCGAGCGATG CCCATATGGG CACGCCCGGA  
5601 TTACAACCCT CCACTGTAG AGTCCTGGAA GGACCCGGAC TACGTCCCTC  
5651 CAGTGGTACA CGGGTGTCCA TTGCCGCTG CCAAGGCCCC TCCGATACCA  
5701 CCTCCACGGA GGAAGAGGAC GGTGTCTCTG TCAGAACTA CCGTGTCTTC  
5751 TGCCTTGGCG GAGCTCGCCA CAAAGACCTT CGGCAGCTCC GAATCGTCGG  
5801 CCGTCGACAG CGGCACGGCA ACGGCCTCTC CTGACCAGCC CTCCGACGAC  
5851 GGCGACGCGG GATCCGACGT TGAGTCGTAC TCCTCCATGC CCCCCCTTGA  
5901 GGGGAGCCG GGGGATCCCG ATCTCAGCGA CGGGTCTTGG TCTACCGTAA  
5951 GCGAGGAGGC TAGTGAGGAC GTCGTCTGCT GCTCGATGTC CTACACATGG  
6001 ACAGGCGCCC TGATCACGCC ATGCGCTGCG GAGGAAACCA AGCTGCCCAT  
6051 CAATGCACTG AGCAACTCTT TGCTCCGTCA CCACAACTTG GTCTATGCTA  
6101 CAACATCTCG CAGCGCAAGC CTGCGGCAGA AGAAGGTCAC CTTTGACAGA  
6151 CTGCAGGTCC TGGACGACCA CTACCGGGAC GTGCTCAAGG AGATGAAGGC

FIG. 1D

6201 GAAGGCGTCC ACAGTTAAGG CTAAACTTCT ATCCGTGGAG GAAGCCTGTA  
6251 AGCTGACGCC CCCACATTCG GCCAGATCTA AATTGGGCTA TGGGGCAAAG  
6301 GACGTCCGGA ACCTATCCAG CAAGGCCGTT AACCACATCC GCTCCGTGTG  
6351 GAAGGACTTG CTGGAAGACA CTGAGACACC AATTGACACC ACCATCATGG  
6401 CAAAAAATGA GGTTTTCTGC GTCCAACCAG AGAAGGGGGG CCGCAAGCCA  
6451 GCTCGCCTTA TCGTATTCCC AGATTTGGGG GTTCGTGTGT GCGAGAAAAT  
6501 GGCCCTTTAC GATGTGGTCT CCACCCTCCC TCAGGCCGTG ATGGGCTCTT  
6551 CATACGGATT CCAATACTCT CCTGGACAGC GGGTCGAGTT CCTGGTGAAT  
6601 GCCTGGAAAG CGAAGAAATG CCCTATGGGC TTCGCATATG ACACCCGCTG  
6651 TTTTGACTCA ACGGTCACTG AGAATGACAT CCGTGTTGAG GAGTCAATCT  
6701 ACCAATGTTG TGA CTGGCC CCCGAAGCCA GACAGGCCAT AAGGTCGCTC  
6751 ACAGAGCGGC TTACATCGG GGGCCCCCTG ACTAATTCTA AAGGGCAGAA  
6801 CTGCGGCTAT CGCCGGTGCC GCGCGAGCGG TGTACTGACG ACCAGCTGCG  
6851 GTAATACCCT CACATGTTAC TTGAAGGCCG CTGCGGCCTG TCGAGCTGCG  
6901 AAGCTCCAGG ACTGCACGAT GCTCGTATGC GGAGACGACC TTGTCGTTAT  
6951 CTGTGAAAGC GCGGGGACCC AAGAGGACGA GGCGAGCCTA CGGGCCTTCA  
7001 CGGAGGCTAT GACTAGATAC TCTGCCCCC CTGGGGACCC GCCCAAACCA  
7051 GAATACGACT TGGAGTTGAT AACATCATGC TCCTCCAATG TGTCAGTCGC  
7101 GCACGATGCA TCTGGCAAAA GGGTGTACTA TCTACCCGT GACCCACCA  
7151 CCCCCCTTGC GCGGGCTGCG TGGGAGACAG CTAGACACAC TCCAGTCAAT  
7201 TCCTGGCTAG GCAACATCAT CATGTATGCG CCCACCTTGT GGGCAAGGAT  
7251 GATCCTGATG ACTCATTTCT TCTCCATCCT TCTAGCTCAG GAACAACTTG  
7301 AAAAAGCCCT AGATTGTCAG ATCTACGGGG CCTGTTACTC CATTGAGCCA  
7351 CTTGACCTAC CTCAGATCAT TCAACGACTC CATGGCCTTA GCGCATTTTC  
7401 ACTCCATAGT TACTCTCCAG GTGAGATCAA TAGGGTGGCT TCATGCCTCA  
7451 GGAAACTTGG GGTACCGCCC TTGCGAGTCT GGAGACATCG GGCCAGAAGT  
7501 GTCCGCGCTA GGCTACTGTC CCAGGGGGGG AGGGCTGCCA CTGTGGCAA  
7551 GTACCTCTTC AACTGGGCAG TAAGGACCAA GCTCAAATC ACTCCAATCC  
7601 CGGCTGCGTC CCAGTTGGAT TTATCCAGCT GGTTCGTTGC TGGTTACAGC  
7651 GGGGGAGACA TATATCACAG CCTGTCTCGT GCGGACCCC GCTGGTTTCA  
7701 GTGGTGCCTA CTCCTACTTT CTGTAGGGGT AGGCATCTAT CTACTCCCCA

FIG. 1E

7751 ACCGATGAAC GGGGAGCTAA ACACTCCAGG CCAATAGGCC ATCCTGTTTT  
7801 TTTCCTTTT TTTTTCCTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT  
7851 TTCTCCTTTT TTTTTCCTCT TTTTTTCCTT TTCTTTCCTT TGGTGGCTCC  
7901 ATCTTAGCCC TAGTCACGGC TAGCTGTGAA AGGTCCGTGA GCCGCTTGAC  
7951 TGCAGAGAGT GCTGATACTG GCCTCTCTGC AGATCAAGTA CTTCTAGAGA  
8001 ATTCTAGCTT GCGTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATGT  
8051 TATCAGCTCA CAATTCACA CAACATACGA GCCGGAAGCA TAAAGTGTA  
8101 AGCCTGGGAT GCCTAATGAG TGAGCTAACT CACATTAGTT GCGTTGCGCT  
8151 CACTGCCCCG TTTCCAGTCG GGAAACCTGT CGTGCCAGCT CCATTAGTGA  
8201 ATCGTCCAAC GCACGGGGAG AGGCGGTTTG CGTATTGGGC GCACTTCCGC  
8251 TTCCTCGCTC ACTGACTCGC TGCCTCGTT CGTTCGGCTG CGGCGAGCCG  
8301 TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT  
8351 AACGCAGGAA AGACCATGTG AGCAAAAGGC CAGCAAAAGG CCAGGAACCG  
8401 TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC CCCCTGACG  
8451 AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA  
8501 CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCCTCTCC  
8551 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG  
8601 GAAGCGTGGC GCTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG  
8651 TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC CCGTTCAGCC  
8701 CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC AACCCGGTAA  
8751 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA  
8801 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA  
8851 CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG  
8901 TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC  
8951 GCTGGTAGCG GTGGTTTTTT TGTGTGCAAG CAGCAGATTA CGCGCAGAAA  
9001 AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC  
9051 AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA  
9101 AGGATCTTCA CCTAGATCCT TTAAATTAA AAATGAAGTT TTAAATCAAT  
9151 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA  
9201 GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCATC CATAGTTGCC  
9251 TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG

FIG. 1F

9301 CCCCAGTGCT GCAATGATAC CGCGAGAACC ACGCTCACCC GCACCAGATT  
9351 TATCAGCAAT AAACCAGCCA GCCGGAAGTG CGCTGCGGAG AAGTGGTCCT  
9401 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AGTTGTTGCC GGGAAGCTAG  
9451 AGTAAGTAGT TCGCCAGTCA GCAGTTTGCG TAACGTCGTT GCCATAGCAA  
9501 CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC  
9551 GGCTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA  
9601 AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG  
9651 CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC  
9701 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC  
9751 ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA  
9801 TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT  
9851 GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG  
9901 ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT  
9951 TTACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC  
10001 GCAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT  
10051 CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG  
10101 GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC  
10151 ACATTTCCCC GAAAAGTGCC ACCTGACGTC TAAGAAACCA TTATTACCAT  
10201 GACATTAACC TATAAAAATA GCGGTATCAC GAAGCCCTTT CGTCTAGCGC  
10251 GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACTTGCAGCT CCCGCAGACG  
10301 GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGGCAAG CCCGTCAGGG  
10351 CGCGTCAGTG GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT  
10401 CAGAGCAGAT TGTACTGAGA GTACACCAGA TGCGGTGTGA AATACGCAC  
10451 AGATGCGTAA GGAGAAAATA CCGCATCAGC CTCCATTCGC CATTCACT  
10501 CCGCAACTGT TGGGAAGGGC GGTCAGTACG CGCTTCTTCG CTATTACGCC  
10551 AACTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA  
10601 GGGTTTTCCC AATCACGACG TTGTAAAACG ACAGCCAATG AATTGAAGCT  
10651 TATTAATTCT AGACTGAAGC TTTAATACG ACTCACTATA (SEQ. ID. NO.:3)

Fig. 1G

IAP12 Rec'd PCT/PTO 01 MAY 2006

## SEQUENCE LISTING

&lt;110&gt; Istituto Di Ricerche Di Biologia Molecolare P. Angeletti S.P.A.

<120> HEPATITIS C VIRUS REPLICONS AND REPLICON  
ENHANCED CELLS

&lt;130&gt; IT0003 PCT

&lt;150&gt; 60/263,479

&lt;151&gt; 2001-01-23

&lt;160&gt; 13

&lt;170&gt; FastSEQ for Windows Version 4.0

&lt;210&gt; 1

&lt;211&gt; 3010

&lt;212&gt; PRT

&lt;213&gt; Con 1 HCV isolate nucleic acid

&lt;400&gt; 1

```

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
1      5      10      15
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
20      25      30
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
35      40      45
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
50      55      60
Ile Pro Lys Ala Arg Gln Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
65      70      75      80
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
85      90      95
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
100     105     110
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
115     120     125
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
130     135     140
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145     150     155     160
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
165     170     175
Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr
180     185     190
Glu Val Arg Asn Val Ser Gly Val Tyr His Val Thr Asn Asp Cys Ser
195     200     205
Asn Ala Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro
210     215     220
Gly Cys Val Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val
225     230     235     240
Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr
245     250     255
Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Ala Ala Ala Leu Cys
260     265     270
Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ala
275     280     285
Gln Leu Phe Thr Phe Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys
290     295     300
Asn Cys Ser Ile Tyr Pro Gly His Val Thr Gly His Arg Met Ala Trp
305     310     315     320

```



```

Asp Met Met Met Asn Trp Ser Pro Thr Ala Ala Leu Val Val Ser Gln
      325      330      335
Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His
      340      345      350
Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp
      355      360      365
Ala Lys Val Leu Ile Val Met Leu Leu Phe Ala Gly Val Asp Gly Gly
      370      375      380
Thr Tyr Val Thr Gly Gly Thr Met Ala Lys Asn Thr Leu Gly Ile Thr
385      390      395      400
Ser Leu Phe Ser Pro Gly Ser Ser Gln Lys Ile Gln Leu Val Asn Thr
      405      410      415
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
      420      425      430
Leu Asn Thr Gly Phe Leu Ala Ala Leu Phe Tyr Val His Lys Phe Asn
      435      440      445
Ser Ser Gly Cys Pro Glu Arg Met Ala Ser Cys Ser Pro Ile Asp Ala
      450      455      460
Phe Ala Gln Gly Trp Gly Pro Ile Thr Tyr Asn Glu Ser His Ser Ser
465      470      475      480
Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile
      485      490      495
Val Pro Ala Ala Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser
      500      505      510
Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Ser
      515      520      525
Trp Gly Glu Asn Glu Thr Asp Val Leu Leu Leu Asn Asn Thr Arg Pro
      530      535      540
Pro Gln Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe
545      550      555      560
Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ile Gly Asn
      565      570      575
Lys Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala
      580      585      590
Thr Tyr Thr Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Leu
      595      600      605
Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe
      610      615      620
Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Leu
625      630      635      640
Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asn Leu Glu Asp
      645      650      655
Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp
      660      665      670
Gln Val Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly
      675      680      685
Leu Ile His Leu His Gln Asn Val Val Asp Val Gln Tyr Leu Tyr Gly
      690      695      700
Ile Gly Ser Ala Val Val Ser Phe Ala Ile Lys Trp Glu Tyr Val Leu
705      710      715      720
Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp
      725      730      735
Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val
      740      745      750
Val Leu Asn Ala Ala Ser Val Ala Gly Ala His Gly Ile Leu Ser Phe
      755      760      765
Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro
      770      775      780
Gly Ala Ala Tyr Ala Leu Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu
785      790      795      800
Leu Ala Leu Pro Pro Arg Ala Tyr Ala Met Asp Arg Glu Met Ala Ala
      805      810      815

```

Ser Cys Gly Gly Ala Val Phe Val Gly Leu Ile Leu Leu Thr Leu Ser  
 820 825 830  
 Pro His Tyr Lys Leu Phe Leu Ala Arg Leu Ile Trp Trp Leu Gln Tyr  
 835 840 845  
 Phe Ile Thr Arg Ala Glu Ala His Leu Gln Val Trp Ile Pro Pro Leu  
 850 855 860  
 Asn Val Arg Gly Gly Arg Asp Ala Val Ile Leu Leu Thr Cys Ala Ile  
 865 870 875 880  
 His Pro Glu Leu Ile Phe Thr Ile Thr Lys Ile Leu Leu Ala Ile Leu  
 885 890 895  
 Gly Pro Leu Met Val Leu Gln Ala Gly Ile Thr Lys Val Pro Tyr Phe  
 900 905 910  
 Val Arg Ala His Gly Leu Ile Arg Ala Cys Met Leu Val Arg Lys Val  
 915 920 925  
 Ala Gly Gly His Tyr Val Gln Met Ala Leu Met Lys Leu Ala Ala Leu  
 930 935 940  
 Thr Gly Thr Tyr Val Tyr Asp His Leu Thr Pro Leu Arg Asp Trp Ala  
 945 950 955 960  
 His Ala Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe  
 965 970 975  
 Ser Asp Met Glu Thr Lys Val Ile Thr Trp Gly Ala Asp Thr Ala Ala  
 980 985 990  
 Cys Gly Asp Ile Ile Leu Gly Leu Pro Val Ser Ala Arg Arg Gly Arg  
 995 1000 1005  
 Glu Ile His Leu Gly Pro Ala Asp Ser Leu Glu Gly Gln Gly Trp Arg  
 1010 1015 1020  
 Leu Leu Ala Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu Leu  
 1025 1030 1035 1040  
 Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Arg Asn Gln Val Glu  
 1045 1050 1055  
 Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr  
 1060 1065 1070  
 Cys Val Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Ser Lys  
 1075 1080 1085  
 Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val  
 1090 1095 1100  
 Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Leu  
 1105 1110 1115 1120  
 Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His  
 1125 1130 1135  
 Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu  
 1140 1145 1150  
 Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro  
 1155 1160 1165  
 Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala Val  
 1170 1175 1180  
 Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu Ser  
 1185 1190 1195 1200  
 Met Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro  
 1205 1210 1215  
 Pro Ala Val Pro Gln Thr Phe Gln Val Ala His Leu His Ala Pro Thr  
 1220 1225 1230  
 Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly  
 1235 1240 1245  
 Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe  
 1250 1255 1260  
 Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr  
 1265 1270 1275 1280  
 Gly Val Arg Thr Ile Thr Thr Gly Ala Pro Ile Thr Tyr Ser Thr Tyr  
 1285 1290 1295  
 Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile  
 1300 1305 1310

Ile Ile Cys Asp Glu Cys His Ser Thr Asp Ser Thr Thr Ile Leu Gly  
 1315 1320 1325  
 Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val  
 1330 1335 1340  
 Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro  
 1345 1350 1355 1360  
 Asn Ile Glu Glu Val Ala Leu Ser Ser Thr Gly Glu Ile Pro Phe Tyr  
 1365 1370 1375  
 Gly Lys Ala Ile Pro Ile Glu Thr Ile Lys Gly Gly Arg His Leu Ile  
 1380 1385 1390  
 Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Ser  
 1395 1400 1405  
 Gly Leu Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser  
 1410 1415 1420  
 Val Ile Pro Thr Ser Gly Asp Val Ile Val Val Ala Thr Asp Ala Leu  
 1425 1430 1435 1440  
 Met Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr  
 1445 1450 1455  
 Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile  
 1460 1465 1470  
 Glu Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg  
 1475 1480 1485  
 Gly Arg Thr Gly Arg Gly Arg Met Gly Ile Tyr Arg Phe Val Thr Pro  
 1490 1495 1500  
 Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys  
 1505 1510 1515 1520  
 Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser  
 1525 1530 1535  
 Val Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln  
 1540 1545 1550  
 Asp His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile  
 1555 1560 1565  
 Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro  
 1570 1575 1580  
 Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro  
 1585 1590 1595 1600  
 Pro Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro  
 1605 1610 1615  
 Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln  
 1620 1625 1630  
 Asn Glu Val Thr Thr Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys  
 1635 1640 1645  
 Met Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly  
 1650 1655 1660  
 Gly Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Thr Thr Gly Ser Val  
 1665 1670 1675 1680  
 Val Ile Val Gly Arg Ile Ile Leu Ser Gly Lys Pro Ala Ile Ile Pro  
 1685 1690 1695  
 Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys Ala  
 1700 1705 1710  
 Ser His Leu Pro Tyr Ile Glu Gln Gly Met Gln Leu Ala Glu Gln Phe  
 1715 1720 1725  
 Lys Gln Lys Ala Ile Gly Leu Leu Gln Thr Ala Thr Lys Gln Ala Glu  
 1730 1735 1740  
 Ala Ala Ala Pro Val Val Glu Ser Lys Trp Arg Thr Leu Glu Ala Phe  
 1745 1750 1755 1760  
 Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala  
 1765 1770 1775  
 Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala  
 1780 1785 1790  
 Phe Thr Ala Ser Ile Thr Ser Pro Leu Thr Thr Gln His Thr Leu Leu  
 1795 1800 1805

Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Pro Pro Ser  
 1810 1815 1820  
 Ala Ala Ser Ala Phe Val Gly Ala Gly Ile Ala Gly Ala Ala Val Gly  
 1825 1830 1835 1840  
 Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly  
 1845 1850 1855  
 Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Val Met Ser Gly Glu  
 1860 1865 1870  
 Met Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser  
 1875 1880 1885  
 Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg  
 1890 1895 1900  
 His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile  
 1905 1910 1915 1920  
 Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro  
 1925 1930 1935  
 Glu Ser Asp Ala Ala Ala Arg Val Thr Gln Ile Leu Ser Ser Leu Thr  
 1940 1945 1950  
 Ile Thr Gln Leu Leu Lys Arg Leu His Gln Trp Ile Asn Glu Asp Cys  
 1955 1960 1965  
 Ser Thr Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Ile  
 1970 1975 1980  
 Cys Thr Val Leu Thr Asp Phe Lys Thr Trp Leu Gln Ser Lys Leu Leu  
 1985 1990 1995 2000  
 Pro Arg Leu Pro Gly Val Pro Phe Phe Ser Cys Gln Arg Gly Tyr Lys  
 2005 2010 2015  
 Gly Val Trp Arg Gly Asp Gly Ile Met Gln Thr Thr Cys Pro Cys Gly  
 2020 2025 2030  
 Ala Gln Ile Thr Gly His Val Lys Asn Gly Ser Met Arg Ile Val Gly  
 2035 2040 2045  
 Pro Arg Thr Cys Ser Asn Thr Trp His Gly Thr Phe Pro Ile Asn Ala  
 2050 2055 2060  
 Tyr Thr Thr Gly Pro Cys Thr Pro Ser Pro Ala Pro Asn Tyr Ser Arg  
 2065 2070 2075 2080  
 Ala Leu Trp Arg Val Ala Ala Glu Glu Tyr Val Glu Val Thr Arg Val  
 2085 2090 2095  
 Gly Asp Phe His Tyr Val Thr Gly Met Thr Thr Asp Asn Val Lys Cys  
 2100 2105 2110  
 Pro Cys Gln Val Pro Ala Pro Glu Phe Phe Thr Glu Val Asp Gly Val  
 2115 2120 2125  
 Arg Leu His Arg Tyr Ala Pro Ala Cys Lys Pro Leu Leu Arg Glu Glu  
 2130 2135 2140  
 Val Thr Phe Leu Val Gly Leu Asn Gln Tyr Leu Val Gly Ser Gln Leu  
 2145 2150 2155 2160  
 Pro Cys Glu Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr  
 2165 2170 2175  
 Asp Pro Ser His Ile Thr Ala Glu Thr Ala Lys Arg Arg Leu Ala Arg  
 2180 2185 2190  
 Gly Ser Pro Pro Ser Leu Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala  
 2195 2200 2205  
 Pro Ser Leu Lys Ala Thr Cys Thr Thr Arg His Asp Ser Pro Asp Ala  
 2210 2215 2220  
 Asp Leu Ile Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn  
 2225 2230 2235 2240  
 Ile Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe  
 2245 2250 2255  
 Glu Pro Leu Gln Ala Glu Glu Asp Glu Arg Glu Val Ser Val Pro Ala  
 2260 2265 2270  
 Glu Ile Leu Arg Arg Ser Arg Lys Phe Pro Arg Ala Met Pro Ile Trp  
 2275 2280 2285  
 Ala Arg Pro Asp Tyr Asn Pro Pro Leu Leu Glu Ser Trp Lys Asp Pro  
 2290 2295 2300

Asp Tyr Val Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Ala Lys  
 2305 2310 2315 2320  
 Ala Pro Pro Ile Pro Pro Pro Arg Arg Lys Arg Thr Val Val Leu Ser  
 2325 2330 2335  
 Glu Ser Thr Val Ser Ser Ala Leu Ala Glu Leu Ala Thr Lys Thr Phe  
 2340 2345 2350  
 Gly Ser Ser Glu Ser Ser Ala Val Asp Ser Gly Thr Ala Thr Ala Ser  
 2355 2360 2365  
 Pro Asp Gln Pro Ser Asp Asp Gly Asp Ala Gly Ser Asp Val Glu Ser  
 2370 2375 2380  
 Tyr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu  
 2385 2390 2395 2400  
 Ser Asp Gly Ser Trp Ser Thr Val Ser Glu Glu Ala Ser Glu Asp Val  
 2405 2410 2415  
 Val Cys Cys Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro  
 2420 2425 2430  
 Cys Ala Ala Glu Glu Thr Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser  
 2435 2440 2445  
 Leu Leu Arg His His Asn Leu Val Tyr Ala Thr Thr Ser Arg Ser Ala  
 2450 2455 2460  
 Ser Leu Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp  
 2465 2470 2475 2480  
 Asp His Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr  
 2485 2490 2495  
 Val Lys Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro  
 2500 2505 2510  
 Pro His Ser Ala Arg Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg  
 2515 2520 2525  
 Asn Leu Ser Ser Lys Ala Val Asn His Ile Arg Ser Val Trp Lys Asp  
 2530 2535 2540  
 Leu Leu Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys  
 2545 2550 2555 2560  
 Asn Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala  
 2565 2570 2575  
 Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met  
 2580 2585 2590  
 Ala Leu Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser  
 2595 2600 2605  
 Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val  
 2610 2615 2620  
 Asn Ala Trp Lys Ala Lys Lys Cys Pro Met Gly Phe Ala Tyr Asp Thr  
 2625 2630 2635 2640  
 Arg Cys Phe Asp Ser Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu  
 2645 2650 2655  
 Ser Ile Tyr Gln Cys Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile  
 2660 2665 2670  
 Arg Ser Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser  
 2675 2680 2685  
 Lys Gly Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu  
 2690 2695 2700  
 Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ala Ala  
 2705 2710 2715 2720  
 Ala Cys Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly  
 2725 2730 2735  
 Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Glu  
 2740 2745 2750  
 Ala Ser Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro  
 2755 2760 2765  
 Pro Gly Asp Pro Pro Lys Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser  
 2770 2775 2780  
 Cys Ser Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val  
 2785 2790 2795 2800

```
<210> 2
<211> 9605
<212> DNA
<213> Con 1 HCV isolate amino acid
```

|             |             |             |             |            |             |      |
|-------------|-------------|-------------|-------------|------------|-------------|------|
| gccagccccc  | gattgggggc  | gacactccac  | catagatcac  | tcccctgtga | ggaactactg  | 60   |
| tcttcacgca  | gaaagcgtct  | agccatggcg  | ttagtatgag  | tgtcgtgcag | cctccaggac  | 120  |
| ccccctccc   | gggagagcca  | tagtggctctg | cggaaccggt  | gagtacaccg | gaattgccag  | 180  |
| gacgacggg   | ctctttcttg  | gatcaaccgc  | ctcaatcgct  | ggagatttgg | gcgtgcccc   | 240  |
| gcgagactgc  | tagccgagta  | gtgttgggtc  | gcgaaaggcc  | tgttggtact | gcctgatagg  | 300  |
| gtgcttgcga  | gtgccccggg  | aggtctcgta  | gaccgtgcac  | catgagcacg | aatcctaaac  | 360  |
| ctcaaaagaaa | aaccaaactg  | aacaccaacc  | gccgccaca   | ggagctcaag | ttcccggggc  | 420  |
| gtggctcagat | cgctgggtgga | gtttacctgt  | tggcgcgcag  | agggcccagg | tggggtgtgc  | 480  |
| gcgcgactag  | gaagactttcc | gagcggctgc  | aacctcgtgg  | aaggcgacaa | cctatcccca  | 540  |
| aggctcgcca  | gcccagaggt  | agggcctggg  | ctcagcccgg  | gtacccctgg | ccccctctatg | 600  |
| gcaatgaggg  | cttgggggtg  | gcaggatggc  | tcctgtcacc  | ccgtggctct | cggcctagtt  | 660  |
| ggggccccc   | ggacccccgg  | cgtaggtcgc  | gcaatttggg  | taaggtcatc | gataccctca  | 720  |
| cgtgcggctt  | gcgcgacttc  | atggggtaca  | ttcgcctcgt  | cgagccccc  | ctaggggggc  | 780  |
| ctgccagggc  | cctggcgcac  | ggcgtccggg  | ttctggagga  | cggcgtgaac | tatgcaacag  | 840  |
| ggaatctgcc  | cggtttgctc  | ttttctatct  | tccttttggc  | tttgctgtcc | tgtttgacca  | 900  |
| tcccagcttc  | cgcttatgaa  | gtgcgcaacg  | tatccggagt  | gtaccatgtc | acgaacgact  | 960  |
| gctcaacgc   | aagcattgtg  | tatgagcgag  | cggacatgat  | catgcatacc | cccgggtgcg  | 1020 |
| tgccctgcgt  | tcggggagaa  | aactctcccc  | gctgctgggt  | agcgtcact  | cccacgtctg  | 1080 |
| cggccaggaa  | cgctagcgtc  | cccactacga  | cgatacgacg  | ccatgtcgat | ttgctcgttg  | 1140 |
| ggggggcctgc | tctctgtccc  | gctatgtacg  | tgggagatct  | ctgcggatct | gttttccctcg | 1200 |
| tcgcccagct  | gtttcaccttc | tcgcctcgcc  | ggcacagacg  | agtcacaggc | tgcaattgtct | 1260 |
| caatatatcc  | cggccacgtg  | acagggtcac  | gtatggcttg  | ggatatgatg | atgaactgggt | 1320 |
| cacctacagc  | agccctagtg  | gtatcgcagt  | tactccggat  | cccacaagct | gtcgtgggata | 1380 |
| tggtagggcgg | ggcccatatg  | ggagctcctag | cgggccttgc  | ctactattcc | atggtggggga | 1440 |
| actgggctaa  | ggttcttgatt | gtgatgtctac | tccttgccgg  | cgttgacggg | ggaaccttatg | 1500 |
| tgacaggggg  | gacgatggcc  | aaaaaacacc  | tcgggatctac | gtccctcttt | tcacccggtg  | 1560 |
| catcccagaa  | aatccagctt  | gtaaacacca  | acggcagctg  | gcacatcaac | aggactggcc  | 1620 |

|             |            |            |             |             |            |      |
|-------------|------------|------------|-------------|-------------|------------|------|
| tgaactgcaa  | tgactccctc | aacactgggt | tccttgctgc  | gctgttctac  | gtgcacaagt | 1680 |
| tcaactcatc  | tggatgcccc | gagcgcatgg | ccagctgcag  | ccccatcgac  | gcgttcgctc | 1740 |
| aggggtgggg  | gcccatact  | tacaatgagt | cacacagctc  | ggaccagagg  | ccttattggt | 1800 |
| ggcactacgc  | accccgccg  | tgcggtatcg | taccgcggc   | gcaggtgtgt  | ggtccagtgt | 1860 |
| actgttcac   | cccaagccct | gtcgtggtgg | ggacgaccga  | cgggttcggc  | gtccctacgt | 1920 |
| acagttgggg  | ggagaatgag | acggacgtgc | tgcttcttaa  | caacacgcgg  | ccgccgcaag | 1980 |
| gcaactgggt  | tggctgtaca | tggatgaata | gcactgggtt  | caccaagacg  | tgcgggggcc | 2040 |
| ccccgtgtaa  | catcgggggg | atcggcaata | aaaccttgac  | ctgccccacg  | gactgtctcc | 2100 |
| ggaagcaccc  | cgaggccact | tacaccaagt | gtgggtcggg  | gccttggttg  | acaccagat  | 2160 |
| gcttggtcca  | ctaccatac  | aggctttggc | actacccctg  | cactgtcaac  | tttaccatct | 2220 |
| tcaaggtag   | gatgtacgtg | gggggagtg  | agcacaggct  | cgaagccgca  | tgcaattgga | 2280 |
| ctcgaggaga  | gcgttgtaac | ctggaggaca | gggacagatc  | agagcttagc  | ccgctgctgc | 2340 |
| tgtctacaac  | ggagtggcag | gtattgccct | gttccttcac  | caccctaccg  | gctctgtcca | 2400 |
| ctggtttgat  | ccatctccat | cagaacgtcg | tggacgtaca  | atacctgtac  | ggtatagggt | 2460 |
| cggcggttgt  | ctcctttgca | atcaaattgg | agtatgtcct  | gttgctcttc  | cttcttctgg | 2520 |
| cggacgcgcg  | cgtctgtgcc | tgcttgtgga | tgatgctgct  | gatagctcaa  | gctgaggccg | 2580 |
| ccctagagaa  | cctggtggtc | ctcaacgcgg | catccgtggc  | cggggcgcat  | ggcattctct | 2640 |
| ccttctctgt  | gttcttctgt | gctgcctggt | acatcaagg   | caggctggtc  | cctggggcgg | 2700 |
| catatgccct  | ctacggcgta | tggccgtac  | tctgtcctc   | gctggcggtt  | ccaccacgag | 2760 |
| catacgccat  | ggaccgggag | atggcagcat | cggtcggagg  | cgcggttttc  | gtaggctga  | 2820 |
| tactcttgac  | cttgtcaccg | cactataagc | tgctcctcgc  | taggctcata  | tggtggttac | 2880 |
| aatattttat  | caccagggcc | gaggcacact | tgcaagtgtg  | gatccccccc  | ctcaacgttc | 2940 |
| ggggggggccg | cgatgcgctg | atcctctca  | cgctgcgcgt  | ccaccagag   | ctaattctta | 3000 |
| ccatcaccaa  | aatcttgctc | gccatactcg | gtccactcat  | ggtgctccag  | gctggtataa | 3060 |
| ccaaagtgcc  | gtacttcgtg | cgcgcacacg | ggctcattcg  | tgcatgcatg  | ctggtgcgga | 3120 |
| aggttgctgg  | gggtcattat | gtccaaatgg | ctctcatgaa  | gttgcccgca  | ctgacaggta | 3180 |
| cgtacgttta  | tgaccatctc | acccactgc  | gggactgggc  | ccacgcgggc  | ctacgagacc | 3240 |
| ttgcggtggc  | agttagcccc | gtcgtcttct | ctgatattga  | gaccaagggt  | atcacctggg | 3300 |
| gggcagacac  | cgcgcggtgt | ggggacatca | tcttgggccc  | gcccgtctcc  | gcccgcaggg | 3360 |
| ggagggagat  | acatctggga | ccggcagaca | gccttgaagg  | gcaggggtgg  | cgactcctcg | 3420 |
| cgctattac   | ggcctactcc | caacagacgc | gaggcctact  | tggtctgcatc | atcactagcc | 3480 |
| tcacaggccg  | ggacaggaac | caggtcgagg | gggagttcca  | agtggtctcc  | accgcaacac | 3540 |
| aatctttcct  | ggcgacctgc | gtcaatggcg | tgtgttggac  | tgtctatcat  | ggtgccggct | 3600 |
| caaagaccct  | tgccggccca | aaggggccaa | tcacccaaat  | gtacaccaat  | gtggaccagg | 3660 |
| acctcgctcg  | ctggcaagcg | cccccgggg  | cgcgctcctt  | gacaccatgc  | acctgcggca | 3720 |
| gctcggaact  | ttacttggtc | acgaggcatg | ccgatgtcat  | tccggtgcgc  | cggcggggcg | 3780 |
| acagcagggg  | gagcctactc | tccccaggc  | ccgtctccta  | cttgaagggc  | tcttcgggcg | 3840 |
| gtccactgct  | ctgccccctg | gggcacgctg | tgggcatctt  | tccggctgcc  | gtgtgcaccc | 3900 |
| gagggggtgc  | gaaggcggtg | gactttgtac | ccgtcgagtc  | tatggaaacc  | actatgcggg | 3960 |
| ccccggtctt  | cacggacaac | tcgtcccctc | cgcccgctacc | gcagacattc  | caggtggccc | 4020 |
| atctacacgc  | ccctactggt | agcggcaaga | gcactaaggt  | gccggctgcg  | tatgcagccc | 4080 |
| aagggtataa  | ggtgcttgct | ctgaacccgt | ccgtcgccgc  | caccctaggt  | ttcggggcgt | 4140 |
| atatgtctaa  | ggcacatggt | atcgacccta | acatcagaac  | cggggtaagg  | accatcacca | 4200 |
| cgggtgcccc  | catcacgtac | tccacctatg | gcaagtttct  | tgccgacggt  | ggttgctctg | 4260 |
| ggggcgccca  | tgacatcata | atatgtgatg | agtgccactc  | aactgactcg  | accactatcc | 4320 |
| tgggcatcgg  | cacagtcctg | gaccaagcgg | agacggctgg  | agcgcgactc  | gtcgtgctcg | 4380 |
| ccaccgctac  | gcctccggga | tcggtcaccg | tgccacatcc  | aaacatcgag  | gaggtggctc | 4440 |
| tgtccagcac  | tggagaaatc | cccttttatg | gcaaagccat  | ccccatcgag  | accatcaagg | 4500 |
| gggggaggca  | cctcattttc | tgccattcca | agaagaaatg  | tgatgagctc  | gccgcgaagc | 4560 |
| tgtccggcct  | cggactcaat | gctgtagcat | attaccgggg  | ccttgatgta  | tccgtcatac | 4620 |
| caactagcgg  | agacgtcatt | gtcgtagcaa | cggacgctct  | aatgacgggc  | tttaccggcg | 4680 |
| atttcgactc  | agtgatcgac | tgcaatacat | gtgtcaccca  | gacagtcgac  | ttcagcctgg | 4740 |
| acccgacctt  | caccattgag | acgacgaccg | tgccacaaga  | cgcggtgtca  | cgctcgacgc | 4800 |
| ggcgaggcag  | gactggtagg | ggcaggatgg | gcatttacag  | gtttgtgact  | ccaggagaac | 4860 |
| ggccctcggg  | catgttcgat | tcctcggttc | tgtgcgagtg  | ctatgacgcy  | ggctgtgctt | 4920 |
| ggtacgagct  | cacgcccgc  | gagacctcag | ttaggttgcy  | ggcttaccta  | aacacaccag | 4980 |
| ggttgcccg   | ctgccaggac | catctggagt | tctgggagag  | cgtctttaca  | ggcctcaccc | 5040 |
| acatagacgc  | ccatttcttg | tcccagacta | agcaggcagg  | agacaacttc  | ccctacctgg | 5100 |
| tagcatacca  | ggctacgggt | tgcgcccagg | ctcaggctcc  | acctccatcg  | tgggaccaaa | 5160 |
| tgtggaagtg  | tctcatacgg | ctaaagccta | cgctgcacgg  | gccaacgccc  | ctgctgtata | 5220 |
| ggtcgggagc  | cgttcaaaac | gaggttacta | ccacacaccc  | cataaccaaa  | tacatcatgg | 5280 |
| catgcatgtc  | ggctgacctg | gaggtcgtca | cgagcacctg  | ggtgctggta  | ggcggagtc  | 5340 |

|             |            |             |             |             |             |      |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| tagcagctct  | ggccgcgtat | tgcctgacaa  | caggcagcgt  | ggtcattgtg  | ggcaggatca  | 5400 |
| tcttgtccgg  | aaagccggcc | atcattcccc  | acagggaagt  | cctttaccgg  | gagttcgatg  | 5460 |
| agatggaaga  | gtgcgcctca | cacctccctt  | acatcgaaca  | gggaatgcag  | ctcgccgaac  | 5520 |
| aattcaaaca  | gaaggcaatc | gggttgctgc  | aaacagccac  | caagcaagcg  | gaggctgctg  | 5580 |
| ctcccgtggt  | ggaatccaag | tggcggaccc  | tccaagcctt  | ctgggcgaag  | catatgtgga  | 5640 |
| atttcatcag  | cgggatacaa | tatttagcag  | gcttgtccac  | tctgcctggc  | aaccccgcca  | 5700 |
| tagcatcact  | gatggcattc | acagcctcta  | tcaccagccc  | gctcaccacc  | caacataccc  | 5760 |
| tcctgtttaa  | catcctgggg | ggatgggtgg  | ccgcccact   | tgctcctccc  | agcgtgctt   | 5820 |
| ctgctttcgt  | aggcgccggc | atcgctggag  | cggctgttgg  | cagcataggc  | cttgggaagg  | 5880 |
| tgcttggtga  | tattttggca | ggttatggag  | caggggtggc  | aggcgcgctc  | gtggccttta  | 5940 |
| aggtcatatg  | cggcgagatg | ccctccaccg  | aggacctggt  | taacctactc  | cctgctatcc  | 6000 |
| tctcccctgg  | cgccctagtc | gtcggggtcg  | tgtgcgcagc  | gatactgctg  | cggcacgtgg  | 6060 |
| gcccaggggga | gggggctgtg | cagtggatga  | accggctgat  | agcgttcgct  | tcgcggggta  | 6120 |
| accacgtctc  | ccccacgcac | tatgtgcctg  | agagcgacgc  | tcagcacgtg  | gtcactcaga  | 6180 |
| tcctctctag  | tcttaccatc | actcagctgc  | tgaagaggct  | tcaccagtgg  | atcaacgagg  | 6240 |
| actgctccac  | gccatgctcc | ggctcggtgc  | taagagatgt  | ttgggattgg  | atatgcacgg  | 6300 |
| tggtgactga  | tttcaagacc | tggctccagt  | ccaagctcct  | gccgcgattg  | ccgggagtcc  | 6360 |
| ccttcttctc  | atgtcaacgt | gggtacaagg  | gagtcgtggc  | gggcgcaggg  | atcatgcaaa  | 6420 |
| ccacgtcgcc  | atgtggagca | cagatcaccg  | gacatgtgaa  | aaacggttcc  | atgaggatcg  | 6480 |
| tggggcctag  | gacctgtagt | aacacgtggc  | attggaacatt | ccccattaac  | gcgtacacca  | 6540 |
| cgggccccctg | cacgcctccc | ccggcgccaa  | attattctag  | ggcgctgtgg  | cgggtggctg  | 6600 |
| ctgaggagta  | cgtggagggt | acgcgggtgg  | gggatttcca  | ctacgtgacg  | ggcatgacca  | 6660 |
| ctgacaacgt  | aaagtgcctg | tgctcaggtc  | cggcccccca  | attcttcaca  | gaagtggatg  | 6720 |
| gggtgcggtt  | gcacaggtac | gctccagcgt  | gcaaacccct  | cctacgggag  | gaggtcacat  | 6780 |
| tcctggtcgg  | gctcaatcaa | tacctggttg  | ggtcacagct  | cccatgagcg  | cccgaaccgg  | 6840 |
| acgtagcagt  | gctcacttcc | atgctcaccg  | acccctccca  | cattacggcg  | gagacggcta  | 6900 |
| agcgtaggct  | ggccagggga | tctccccctt  | ccttgggcag  | ctcatcagct  | agccagctgt  | 6960 |
| ctgcgccttc  | cttgaaggca | acatgcacta  | ccctgcatga  | ctccccggac  | gctgacctca  | 7020 |
| tcgaggccaa  | cctcctgtgg | cggcaggaga  | tggcggggaa  | catcaccggc  | gtggagttag  | 7080 |
| aaaataaggt  | agtaattttg | gactctttcg  | agccgctcca  | agcggaggag  | gatgagaggg  | 7140 |
| aagtatccgt  | tccggcggag | atcctgcgga  | gggtccaggaa | attccctcga  | gcgatgccca  | 7200 |
| tatggcgctg  | cccgaggtac | aacctccac   | tgtagagtc   | ctggaaggac  | ccggactacg  | 7260 |
| tcctccaggt  | ggtagacggg | tgctccattg  | cgcctgccaa  | ggccccctcg  | ataccacctc  | 7320 |
| cacggaggaa  | gaggacggtt | gtcctgtcag  | aatctaccgt  | gtcttctgcc  | ttggcggagc  | 7380 |
| tcgcccacaa  | gaccttcggc | agctccgaat  | cgtcggccgt  | cgacagcggc  | acggcaacgg  | 7440 |
| cctctgcctc  | ccagccctcc | gacgacggcg  | acgcgggatc  | cgacgttgag  | tcgtactcct  | 7500 |
| ccatgcccc   | ccttgagggg | gagccggggg  | atcccgatct  | cagcgacggg  | tcttggtcta  | 7560 |
| ccgtaagcga  | ggaggctagt | gaggacgtcg  | tctgtgctc   | gatgtcctac  | acatggacag  | 7620 |
| gcgccctgat  | cacgccatgc | gctgcggagg  | aaaccaagct  | gccccatcaat | gcactgagca  | 7680 |
| actctttgct  | ccgtcaccac | aacttggtct  | atgctacaac  | atctcgcagc  | gcaagcctgc  | 7740 |
| ggcagaagaa  | ggcacctttt | gacagactgc  | aggtcctgga  | cgaccactac  | cgggacgtgc  | 7800 |
| tcaaggagat  | gaaggcgaag | gcgtccacag  | ttaaggctaa  | acttctatcc  | gtggagggaag | 7860 |
| cctgtaagct  | gacgccccca | cattcggcca  | gatctaaatt  | tggctatggg  | gcaaaggacg  | 7920 |
| tccggaacct  | atccagcaag | gccgttaacc  | acatccgctc  | cgtgtggaag  | gacttgcctg  | 7980 |
| aagacactga  | gacaccaatt | gacaccacca  | tcatggcaaa  | aaatgagggt  | ttctgcgtcc  | 8040 |
| aaccagagaa  | ggggggccgc | aagccagctc  | gccttatcgt  | attcccagat  | ttgggggttc  | 8100 |
| gtgtgtgcga  | gaaaatggcc | ctttacgatg  | tggtctccac  | cctccctcag  | gccgtgatgg  | 8160 |
| gctcttcata  | cggattccaa | tactctcctg  | gacagcgggt  | cgagtccctg  | gtgaatgcct  | 8220 |
| ggaaagcgaa  | gaaatgccct | atgggcttcg  | catatgacac  | ccgctgtttt  | gactcaacgg  | 8280 |
| tcactgagaa  | tgacatccgt | gttgaggagt  | caatctacca  | atgttgtagc  | ttggcccccg  | 8340 |
| aagccagaca  | ggccataagg | tcgctcacag  | agcggcttta  | catcgggggc  | cccctgacta  | 8400 |
| attctaaagg  | gcagaactgc | ggctatcgcc  | gggtgcgcgc  | gagcgggtga  | ctgacgacca  | 8460 |
| gctgcggtaa  | tacctcaca  | tgttacttga  | aggccgctgc  | ggcctgtcga  | gctgcgaagc  | 8520 |
| tccaggactg  | cacgatgctc | gtatgcggag  | acgaccttgt  | cgttatctgt  | gaaagcgcgg  | 8580 |
| ggacccaaga  | ggacgaggcg | agcctacggg  | ccttcacgga  | ggctatgact  | agatactctg  | 8640 |
| ccccccctgg  | ggacccgccc | aaaccagaat  | acgacttgga  | gttgataaca  | tcagtctcct  | 8700 |
| ccaatgtgtc  | agtgcgcgac | gatgcactctg | gcaaaagggt  | gtactatctc  | acccgtgacc  | 8760 |
| ccaccacccc  | ccttgccggg | gctgcgtggg  | agacagctag  | acacactcca  | gtcaattcct  | 8820 |
| ggctaggcaa  | catcatcatg | tatgcgccca  | ccttggtggc  | aaggatgatc  | ctgatgactc  | 8880 |
| atttcttctc  | catccttcta | gctcagggaac | aacttgaaaa  | agccctagat  | tgtagatctc  | 8940 |
| acggggcctg  | ttactccatt | gagccacttg  | acctacctca  | gatcatccaa  | cgactccatg  | 9000 |
| gccttagcgc  | attttcactc | catagttact  | ctccagggtga | gatcaatagg  | gtggccttcat | 9060 |



|            |            |            |             |            |            |      |
|------------|------------|------------|-------------|------------|------------|------|
| gcctcaggaa | acttggggta | ccgcccttgc | gagtcctggag | acatcggggc | agaagtgtcc | 9120 |
| gcgctaggct | actgtccag  | ggggggaggg | ctgccacttg  | tggcaagtac | ctcttcaact | 9180 |
| gggcagtaag | gaccaagctc | aaactcactc | caatcccggc  | tgcgtcccag | ttggatttat | 9240 |
| ccagctgggt | cgttgctggt | tacagcgggg | gagacatata  | tcacagcctg | tctcgtgccc | 9300 |
| gaccccgtg  | gttcatgtgg | tgccactacc | tactttctgt  | aggggtaggc | atctatctac | 9360 |
| tccccaaccg | atgaacgggg | agctaaacac | tccaggccaa  | tagggcatcc | tgtttttttc | 9420 |
| cctttttttt | tttctttttt | tttttttttt | tttttttttt  | ttttttttct | cctttttttt | 9480 |
| tcctcttttt | ttcctttttt | ttcctttggt | ggctccatct  | tagccctagt | cacggctagc | 9540 |
| tgtgaaagg  | ccgtgagccg | cttgactgca | gagagtgtctg | atactggcct | ctctgcagat | 9600 |
| caagt      |            |            |             |            |            | 9605 |

&lt;210&gt; 3

&lt;211&gt; 10690

&lt;212&gt; DNA

&lt;213&gt; pHCVNeo.17 coding

&lt;400&gt; 3

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gccagcccc   | gattgggggc  | gacactccac  | catagatcac  | tccccgtgta  | ggaactactg  | 60   |
| tcttcacgca  | gaaagcgtct  | agccatggcg  | ttagtatgag  | tgtcgtgcag  | cctccaggac  | 120  |
| ccccctccc   | gggagagcca  | tagtggctctg | cggaaccggg  | gagtacaccg  | gaattgccag  | 180  |
| gacgacggg   | gttcatgtgg  | gatcaaccgg  | ctcaatgcct  | ggagatttgg  | gctgcccc    | 240  |
| gcgagactgc  | tagccgagta  | gtgttgggtc  | gcgaaaggcc  | ttgtggtact  | gcctgatagg  | 300  |
| gtgcttgcca  | gtgccccggg  | aggtctcgta  | gaccgtgcac  | catgagcacg  | aatcctaaac  | 360  |
| ctcaaaagaa  | aaccaaaagg  | cgcgccatga  | ttgaacaaga  | tggattgcac  | gcagggtctc  | 420  |
| cgggcgcttg  | ggtggagagg  | ctattcggct  | atgactgggc  | acaacagaca  | atcgggtgct  | 480  |
| ctgatgccgc  | cgtgttccgg  | ctgtcagcgc  | aggggcggcc  | ggttcttttt  | gtcaagaccg  | 540  |
| acctgtccgg  | tgccctgaat  | gaactgcagg  | acgaggcagc  | gcggctatcg  | tggctggcca  | 600  |
| cgacggggcg  | tccttgcgca  | gctgtgctcg  | acgtttgtac  | tgaagcggga  | agggactggc  | 660  |
| tgtatttggg  | cgaagtgccg  | gggcaggatc  | tctgtctatc  | tcaccttgct  | cctgccgaga  | 720  |
| aagtatccat  | catggctgat  | gcaatgcggc  | ggctgcatac  | gcttgatccg  | gctacctgcc  | 780  |
| cattcgacca  | ccaagcgaaa  | catcgcatcg  | agcgagcacg  | tactcggatg  | gaagccggtc  | 840  |
| ttgtcgatca  | ggatgatctg  | gacgaagagc  | atcaggggct  | cgcgccagcc  | gaactgttcg  | 900  |
| ccaggctcaa  | ggcgcgcatg  | cccgaaggcg  | aggatctcgt  | cgtgacccat  | ggcgatgcct  | 960  |
| gcttgccgaa  | tatctgggtg  | gaaaatggcc  | gcttttcttg  | attcatcgac  | tgtggccggc  | 1020 |
| tgggtgtggc  | ggaccgctat  | caggacatag  | cgttggctac  | ccgtgatatt  | gctgaagagc  | 1080 |
| ttggcgccga  | atgggctgac  | cgcttcctcg  | tgttttacgg  | tatcgccgct  | cccgatctgc  | 1140 |
| agcgcatcgc  | cttctatcgc  | cttcttgacg  | agtttcttctg | agtttaaaca  | gaccacaacg  | 1200 |
| gtttccctct  | agcgggatca  | attccgcccc  | tctccctccc  | ccccccctaa  | cgttactggc  | 1260 |
| cgaagccgct  | tgggaataagg | ccggtgtgcg  | tttgtctata  | tggtattttc  | caccatattg  | 1320 |
| ccgtcttttg  | gcaatgtgag  | ggcccggaaa  | cctggccctg  | tcttcttgac  | gagcattcct  | 1380 |
| aggggtcttt  | cccctctcgc  | caaaggaatg  | caagggtctgt | tgaatgtcgt  | gaaggaagca  | 1440 |
| gttccctctg  | aagcttcttg  | aagacaaaca  | acgtctgtag  | cgaccctttg  | caggcagcgg  | 1500 |
| aacccccccac | ctggcgacag  | gtgcctctgc  | ggccaaaagc  | cacgtgtata  | agatacacct  | 1560 |
| gcaaaggcgg  | cacaacccca  | gtgccacggt  | gtgagttgga  | tagttgtgga  | aagagtcaaa  | 1620 |
| tggctctcct  | caagcgtatt  | caacaagggg  | ctgaaggatg  | cccagaagg   | accccatgtg  | 1680 |
| atgggatctg  | atctggggcc  | tcggtgcaca  | tgttttacat  | gtgttttagtc | gagggttaaaa | 1740 |
| aacgtctagg  | ccccccgaac  | cacggggacg  | tggttttcct  | tgaaaaaaca  | cgataatacc  | 1800 |
| atggcgcccta | ttacggccta  | ctcccaacag  | acgcgaggcc  | tacttggctg  | catcatcact  | 1860 |
| agcctcacag  | gccgggacag  | gaaccaggtc  | gagggggagg  | tccaagtgg   | ctccaccgca  | 1920 |
| acacaatctt  | tcctggcgac  | ctgcgtcaat  | ggcgtgtgtt  | ggactgtcta  | tcatggtgcc  | 1980 |
| ggctcaaaaga | cccttgccgg  | cccaaagggc  | ccaatcaccc  | aaatgtacac  | caatgtggac  | 2040 |
| caggacctcg  | tcggctggca  | agcgcccccc  | ggggcgcggt  | ccttgacacc  | atgcacctgc  | 2100 |
| ggcagctcgg  | acctttactt  | ggtcacgagg  | catgccgatg  | tcattccggt  | gcgccggcgg  | 2160 |
| ggcgacagca  | ggggggagcct | acttcccccc  | agggccgtct  | cctacttgaa  | gggctcttcg  | 2220 |
| ggcggtccac  | tgtcttgccc  | ctcggggcac  | gctgtgggca  | tctttcgggc  | tgcctgtgtg  | 2280 |
| acccgagggg  | tgtcgaaagg  | ggtggacttt  | gtaccgctcg  | agtctatgga  | aaccactatg  | 2340 |
| cggctccccgg | tcttcacgga  | caactcgtcc  | cctccggccg  | taccgcagac  | attccaggtg  | 2400 |
| gcccatactac | acgccccctac | tggtagcggc  | aagagcacta  | aggtgccggc  | tgcgtatgca  | 2460 |
| gcccgaagggt | ataagggtgct | tgtcctgaac  | ccgtccgctg  | ccgccaccct  | aggtttcggg  | 2520 |
| gcgtatatgt  | ctaaggcaca  | tggatctgac  | cctaacaatca | gaaccggggg  | aaggaccatc  | 2580 |

|             |             |             |             |             |             |      |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| accacgggtg  | cccccatcac  | gtactccacc  | tatggcaagt  | ttcttgccga  | cggtggttgc  | 2640 |
| tctgggggcg  | cctatgacat  | cataatatgt  | gatgagtgcc  | actcaactga  | ctcgaccact  | 2700 |
| atcctgggca  | tcggcacagt  | cctggaccaaa | gcgagagcgg  | ctggagcgcg  | actcgtcggtg | 2760 |
| ctcgccaccg  | ctacgcctcc  | gggatcggtc  | accgtgccac  | atccaaacat  | cgaggagggtg | 2820 |
| gctctgtcca  | gcactggaga  | aatccccttt  | tatggcaaaag | ccatccccat  | cgagaccatc  | 2880 |
| aaggggggga  | ggcacctcat  | tttctgccat  | tccaagaaga  | aatgtgatga  | gctcgccgcg  | 2940 |
| aagctgtccg  | gcctcggact  | caatgctgta  | gcatattacc  | ggggccttga  | tgtatccgtc  | 3000 |
| ataccaacta  | gcggagacgt  | cattgtcgta  | gcaacggacg  | ctctaattgac | gggctttacc  | 3060 |
| ggcgatttcg  | actcagtgat  | cgactgcaat  | acatgtgtca  | cccagacagt  | cgacttcagc  | 3120 |
| ctggaccgga  | ccttcacat   | tgagacgacg  | accgtgccac  | aagacgcggt  | gtcacgctcg  | 3180 |
| cagcggcgag  | gcaggactgg  | taggggcagg  | atgggcattt  | acaggtttgt  | gactccagga  | 3240 |
| gaacggccct  | cgggcatggt  | cgattcctcg  | gttctgtgcg  | agtgtatga   | cgcgggctgt  | 3300 |
| gcttggtacg  | agctcacgcc  | cgccgagacc  | tcagttaggt  | tgcgggctta  | cctaaacaca  | 3360 |
| ccagggttgc  | ccgtctgcca  | ggaccatctg  | gagttctggg  | agagcgtctt  | tacaggcctc  | 3420 |
| accacatag   | acgcccattt  | cttgtccag   | actaagcagg  | caggagacaa  | cttcccctac  | 3480 |
| ctggtagcat  | accaggctac  | ggtgtgcgcc  | agggctcagg  | ctccacctcc  | atcgtgggac  | 3540 |
| caaagtgtga  | agtgtctcat  | acggctaaag  | cctacgctgc  | acgggccaaac | gccccgtgtg  | 3600 |
| tataggctgg  | gagccgttca  | aaacgagggt  | actaccacac  | accccataac  | caaatacatc  | 3660 |
| atggcatgga  | tgtcggctga  | cctggaggctc | gtgcaaacga  | cctgggtgct  | ggtaggcgga  | 3720 |
| gtcctagcag  | ctctggccgc  | gtattgcctg  | acaacaggca  | gcgtgggtcat | tgtgggcagg  | 3780 |
| atcatcttgt  | ccggaaagcc  | ggccatcatt  | cccagacagg  | aagtccttta  | ccgggagttc  | 3840 |
| gatgagatgg  | aagagtgcgc  | ctcacacctc  | ccttacatcg  | aacagggaat  | gcagctcgcc  | 3900 |
| gaacaattca  | aacagaaggc  | aatcgggttg  | gtgcaaacag  | ccaccaagca  | agcggaggct  | 3960 |
| gctgctcccg  | tgggtggaatc | caagtggcgg  | accctcgaag  | ccttctgggc  | gaagcatatg  | 4020 |
| tgggaatttca | tcagcgggat  | acaatattta  | gcaggcttgt  | ccactctgcc  | tggcaacccc  | 4080 |
| gcgatagcat  | cactgatggc  | attcacagcc  | tctatcacca  | gcccgcctac  | cacccaacat  | 4140 |
| accctcctgt  | ttaacatcct  | ggggggatgg  | gtggccgccc  | aacttgctcc  | tcccagcgct  | 4200 |
| gcttctgctt  | tcgtaggcgc  | cgccatcgct  | ggagcggctg  | ttggcagcat  | aggccttggg  | 4260 |
| aagggtgctt  | tggatatattt | ggcaggttat  | ggagcagggg  | tggcaggcgc  | gctcgtggcc  | 4320 |
| tttaaggtca  | tgagcggcga  | gatgccctcc  | accgaggacc  | tggttaacct  | actccctgct  | 4380 |
| atcctctccc  | ctggcgccct  | agtcgtcggg  | gtcgtgtgcg  | cagcgatact  | gcgtcggcac  | 4440 |
| gtgggcccag  | gggagggggc  | tgtgcagtgg  | ctgaaccggc  | tgatagcggt  | cgcttcgcgg  | 4500 |
| ggtaaccacg  | tctccccac   | gcactatgtg  | cctgagagcg  | acgctgcagc  | acgtgtcact  | 4560 |
| cagatcctct  | ctagtcttac  | catcactcag  | ctgctgaaga  | ggcttcacca  | gtggatcaac  | 4620 |
| gaggactgct  | ccacgccatg  | ctccggctcg  | tggctaagag  | atgtttggga  | ttggatatgc  | 4680 |
| acggtgttga  | ctgatttcaa  | gacctggctc  | cagtcacaagc | tcctgccgcg  | atttgcggga  | 4740 |
| gtccccctct  | tctcatgtca  | acgtgggtac  | aagggagctc  | ggcggggcga  | cggcatcatg  | 4800 |
| caaaccacct  | gcccattgtg  | agcacagatc  | accggacatg  | tgaaaaacgg  | ttccatgagg  | 4860 |
| atcgtggggc  | ctaggacctg  | tagtaacacg  | tggcatggaa  | cattccccat  | taacgcgtac  | 4920 |
| accacggggc  | cctgcacgcc  | ctccccggcg  | ccaaattatt  | ctagggcgct  | gtggcgggtg  | 4980 |
| gctgctgagg  | agtacgtgga  | ggttacgcgg  | gtgggggatt  | tccactacgt  | gacgggcatg  | 5040 |
| accactgaca  | acgtaaagtg  | cccgtgtcag  | gttccggccc  | ccgaattctt  | cacagaagtg  | 5100 |
| gatgggggtg  | ggttgcacag  | gtacgctcca  | gcgtgcaaac  | ccctcctacg  | ggaggaggtc  | 5160 |
| acatttctgg  | tcgggctcaa  | tcaatacctg  | gttgggtcac  | agctcccatg  | cgagcccgaa  | 5220 |
| ccggacgtag  | cagtgtcac   | ttccatgtct  | accgacccct  | cccacattac  | ggcggagacg  | 5280 |
| gctaagcgta  | ggctggccag  | gggatctccc  | ccctccttgg  | ccagctcatc  | agctagccag  | 5340 |
| ctgtctgctg  | cttcccttga  | ggcaacatgc  | actaccgctc  | atgactcccc  | ggacgctgac  | 5400 |
| ctcactgagg  | ccaacctcct  | gtggcggcag  | gagatggggc  | ggaacatcac  | ccgcgtggag  | 5460 |
| tcagaaaata  | aggtagtaat  | tttggactct  | ttcagaccgc  | tccaagcgga  | ggaggatgag  | 5520 |
| agggaaagtat | ccgttcgggc  | ggagatcctg  | cggagggtcca | ggaaattccc  | tcgagcgatg  | 5580 |
| cccatatggg  | cacgcccggg  | ttacaaccct  | ccactgttag  | agtcctggaa  | ggaccgggac  | 5640 |
| tacgtccctc  | cagtggta    | cgggtgtcca  | ttgccgcctg  | ccaaggcccc  | tccgatacca  | 5700 |
| cctccacgga  | ggaagaggac  | ggttgtcctg  | tcaagaatcta | ccgtgtcttc  | tgccttggcg  | 5760 |
| gagctcgcca  | caaagacctt  | cggcagctcc  | gaatcgtcgg  | ccgtcgacag  | cggcacggca  | 5820 |
| acggcctctc  | ctgaccagcc  | ctccgacgac  | ggcgacgcgg  | gatccgacgt  | tgagtcgtac  | 5880 |
| tcctccatgc  | cccccttga   | gggggagccg  | ggggatcccc  | atctcagcga  | cgggtcttgg  | 5940 |
| tctaccgtaa  | gcgaggaggc  | tagtgaggac  | gtcgtctgct  | gctcgaatgc  | ctacacatgg  | 6000 |
| acaggcgccc  | tgatcacgcc  | atgcgctgcg  | gaggaaacca  | agctgcccac  | caatgcaactg | 6060 |
| agcaactctt  | tgctccgtca  | ccacaacttg  | gtctatgcta  | caacatctcg  | cagcgcaagc  | 6120 |
| ctgcggcaga  | agaagggtcac | ctttgacaga  | ctgcagggtcc | tggacgacca  | ctaccgggac  | 6180 |
| gtgctcaagg  | agatgaaggc  | gaaggcgctc  | acagtttaagg | ctaaacttct  | atccgtggag  | 6240 |
| gaagcctgta  | agctgacgcc  | cccacattcg  | gccagatcta  | aatttggcta  | tggggcaaa   | 6300 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gacgtccgga  | acctatccag  | caaggccgtt  | aaccacatcc  | gctccgtgtg  | gaaggacttg  | 6360  |
| ctggaagaca  | ctgagacacc  | aattgacacc  | accatcatgg  | caaaaaatga  | ggttttctgc  | 6420  |
| gtccaaccag  | agaagggggg  | ccgcaagcca  | gctcgcttta  | tcgtattccc  | agatttgggg  | 6480  |
| gttcgtgtgt  | gcgagaaaat  | ggccctttac  | gatgtgggtc  | ccaccctccc  | tcaggccgtg  | 6540  |
| atgggctctt  | catacggatt  | ccaatactct  | cctggacagc  | gggtcaggtt  | cctgggtgaat | 6600  |
| gcctggaaag  | cgaagaaatg  | ccctatgggc  | ttcgcatatg  | acaccgcgtg  | ttttgactca  | 6660  |
| acggtcactg  | agaatgacat  | ccgtgttgag  | gagtcaatct  | accaatgttg  | tgacttggcc  | 6720  |
| cccgaagcca  | gacaggccat  | aaggctcgctc | acagagcggc  | tttacatcgg  | gggccccctg  | 6780  |
| actaattcta  | aagggcagaa  | ctgcggtcat  | cgccgggtgcc | gcgcgagcgg  | tgtactgacg  | 6840  |
| accagctgcg  | gtaataccct  | cacatgttac  | ttgaaggccg  | ctgcggcctg  | tcgagctgcg  | 6900  |
| aagctccagg  | actgcacgat  | gctcgtatgc  | ggagagcacc  | ttgtcgttat  | ctgtgaaagc  | 6960  |
| gcggggaccc  | aagaggacga  | ggcgagccta  | cgggccttca  | cggaggctat  | gactagatac  | 7020  |
| tctgcccccc  | ctggggaccc  | gccc aaacca | gaatacagact | tggagttagat | aacatcatgc  | 7080  |
| tcttccaatg  | tgtcagtcgc  | gcacgatgca  | tctggcaaaa  | gggtgtacta  | tctcaccctg  | 7140  |
| gacccaccca  | cccccttgc   | gcgggctgcg  | tggagagacg  | ctagacacac  | tccagtcaat  | 7200  |
| tcttggctag  | gcaacatcat  | catgtatgcg  | cccaccttgt  | gggcaaggat  | gactcctgatg | 7260  |
| actcatttct  | tctccatcct  | tctagctcag  | gaacaacttg  | aaaaagccct  | agattgtcag  | 7320  |
| atctacgggg  | cctgttactc  | cattgagcca  | cttgacctac  | ctcagatcat  | tcaacgactc  | 7380  |
| catggcctta  | ggcaactttc  | actccatagt  | tactctccag  | gtgagatcaa  | taggggtggc  | 7440  |
| tcatgcctca  | ggaaacttgg  | ggtaccgccc  | ttgcgagtct  | ggagacatcg  | ggccagaagt  | 7500  |
| gtccgcgcta  | ggctactgtc  | ccaggggggg  | agggctgccca | cttgtggcaa  | gtacctcttc  | 7560  |
| aactgggcag  | taaggaccaa  | gctcaaactc  | actccaatcc  | cggctgcgtc  | ccagttggat  | 7620  |
| ttatccagtc  | ggttcgttgc  | tggttacagc  | gggggagaca  | tatatcacag  | cctgtctcgt  | 7680  |
| gcccgaaccc  | gtgtgttcct  | gtggtgccta  | ctctactttt  | ctgtaggggt  | aggcatctat  | 7740  |
| ctactcccca  | accgatgaac  | ggggagctaa  | acactccagg  | ccaataggcc  | atcctgtttt  | 7800  |
| tttccctttt  | tttttttctt  | tttttttttt  | tttttttttt  | tttttttttt  | ttctcctttt  | 7860  |
| tttttctctt  | tttttctctt  | ttcttctctt  | tggtggctcc  | atcttagccc  | tagtcacggc  | 7920  |
| tagctgtgaa  | aggttccgtga | gccgcttgac  | tgcagagagt  | gctgatactg  | gcctctctgc  | 7980  |
| agatcaagta  | cttctagaga  | attctagctt  | ggcgtaatca  | tggtcatagc  | tgtttctctg  | 8040  |
| gtgaaattgt  | tatcagctca  | caattccaca  | caacatacga  | gccggaagca  | taaagtgtaa  | 8100  |
| agcctgggat  | gcctaattgag | tgagctaact  | cacattagtt  | gcgttgcgct  | cactgcccgc  | 8160  |
| tttccagtcg  | ggaaacctgt  | cgtgccagct  | ccatttagta  | atcgtccaac  | gcacggggag  | 8220  |
| aggcgggttg  | cgtattgggc  | gcacttccgc  | ttcctcgctc  | actgactcgc  | tgcgctcgtt  | 8280  |
| cgttcggtcg  | cggcgagccg  | tatcagctca  | ctcaaaggcg  | gtaatacggg  | tatccacaga  | 8340  |
| atcaggggat  | aacgcaggaa  | agaccatgtg  | agcaaaaggc  | cagcaaaagg  | ccaggaaccg  | 8400  |
| taaaaaaggcc | gcgttgctgg  | cgtttttcca  | taggtccgc   | ccccctgacg  | agcatcacia  | 8460  |
| aaatcgacgc  | tcaagtcaga  | ggtggcgaaa  | cccgcagga   | ctataaagat  | accaggcgtt  | 8520  |
| tccccctgga  | agctccctcg  | tgcgctctcc  | tgttccgacc  | ctgccgctta  | ccggatacct  | 8580  |
| gtccgccttt  | ctcccttcgg  | gaagcggtgg  | gctttctcat  | agctcacgct  | gtaggtatct  | 8640  |
| cagttcggtg  | taggtcgttc  | gctccaagct  | gggctgtgtg  | cacgaacccc  | ccgttcagcc  | 8700  |
| cgaccgctgc  | gccttatccg  | gtaactatcg  | tctttagtcc  | aaccgcgtta  | gacacgactt  | 8760  |
| atcgccactg  | gcagcagcca  | ctggtaacag  | gattagcaga  | gcgaggatag  | taggcggtgc  | 8820  |
| tacagagttc  | ttgaagtggg  | ggcctaacta  | cggctacact  | agaaggacag  | tatttggtat  | 8880  |
| ctgcgctctg  | ctgaagccag  | ttaccttcgg  | aaaaagagtt  | ggtagctctt  | gatccggcaa  | 8940  |
| acaaaccacc  | gctggtagcg  | gtggtttttt  | tgtttgcaag  | cagcagatta  | cgcgcagaaa  | 9000  |
| aaaaggatct  | caagaagatc  | ctttgatctt  | ttctacgggg  | tctgacgctc  | agtggaaacga | 9060  |
| aaactcacgt  | taagggatth  | tggtcatgag  | attatcaaaa  | aggatcttca  | cctagatcct  | 9120  |
| tttaaattaa  | aaatgaagtt  | ttaaatcaat  | ctaaagtata  | tatgagtaaa  | cttggtctga  | 9180  |
| cagttaccaaa | tgcttaatca  | gtgaggcacc  | tatctcagcg  | atctgtctat  | ttcgttcatc  | 9240  |
| catagttgcc  | tgactccccg  | tcgtgtagat  | aactacgata  | cgggagggct  | taccatctgg  | 9300  |
| ccccagtgct  | gcaatgatac  | cgcgagaacc  | acgctcacc   | gcaccagatt  | tatcagcaat  | 9360  |
| aaaccagcca  | gccggaagtg  | cgctgcggag  | aagtggctct  | gcaactttat  | ccgcctccat  | 9420  |
| ccagttctatt | agttgtttgc  | gggaagctag  | agtaagtagt  | tcgccagtc   | gcagtttgcc  | 9480  |
| taacgtcggt  | gcatagcaaa  | caggcatcgt  | ggtgtcacgc  | tcgtcgtttg  | gtatggcttc  | 9540  |
| atcagctcc   | ggctcccaac  | gatcaaggcg  | agttacatga  | tccccatgt   | tgtgcaaaaa  | 9600  |
| agcgggttagc | tccttcgggtc | ctccgatcgt  | tgtcagaagt  | aagttggccg  | cagtgttatc  | 9660  |
| actcatgggt  | atggcagcac  | tgcataattc  | tcttactgtc  | atgccatccg  | taagatgctt  | 9720  |
| ttctgtgact  | ggtgtgtact  | caaccaagtc  | attctgagaa  | tagtgtatgc  | ggcgaccgag  | 9780  |
| ttgtctcttc  | ccggcgtaaa  | tacgggataa  | taccgcgcca  | catagcagaa  | ctttaaaggt  | 9840  |
| gctcatcatt  | ggaaaacgtt  | cttcggggcg  | aaaactctca  | aggatcttac  | cgctgttgag  | 9900  |
| atccagttcg  | atgtaaccca  | ctcgtgcacc  | caactgatct  | tcagcatctt  | ttactttcac  | 9960  |
| cagcgtttct  | gggtgagcaa  | aaacaggaag  | gcaaaatgcc  | gcaaaaaagg  | gaataagggc  | 10020 |

|            |            |            |            |             |             |       |
|------------|------------|------------|------------|-------------|-------------|-------|
| gacacggaaa | tggtgaatac | tcatactctt | cctttttcaa | tattattgaa  | gcatttatca  | 10080 |
| gggttattgt | ctcatgagcg | gatacatatt | tgaatgtatt | tagaaaaata  | aacaaatagg  | 10140 |
| ggttccgcgc | acatttcccc | gaaaagtgcc | acctgacgtc | taagaaacca  | ttattaccat  | 10200 |
| gacattaacc | tataaaaata | ggcgtatcac | gaagcccttt | cgtctagcgc  | gtttcgggtga | 10260 |
| tgacggtgaa | aacctctgac | acttgcagct | cccgcagacg | gtcacagctt  | gtctgtaagc  | 10320 |
| ggatgccggg | agcaggcaag | cccgtcaggg | cgcgtcagtg | gggtgttgccg | gggtgtcgggg | 10380 |
| ctggcttaac | tatgcggcat | cagagcagat | tgtactgaga | gtacaccaga  | tgcggtgtga  | 10440 |
| aataccgcac | agatgcgtaa | ggagaaaata | ccgcatcagc | ctccattcgc  | cattcagact  | 10500 |
| ccgcaactgt | tggaagggc  | ggtcagtacg | cgcttcttcg | ctattacgcc  | aactggcgaa  | 10560 |
| aggggatgt  | gctgcaaggc | gattaagttg | ggtaacgcca | gggttttccc  | aatcacgacg  | 10620 |
| ttgtaaacg  | acagccaatg | aattgaagct | tattaattct | agactgaagc  | ttttaatacg  | 10680 |
| actcactata |            |            |            |             |             | 10690 |

&lt;210&gt; 4

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Primer oligonucleotide

&lt;400&gt; 4

acatgatctg cagagaggcc agt

23

&lt;210&gt; 5

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Primer oligonucleotide

&lt;400&gt; 5

gacasgctgt gatawatgtc tcccc

26

&lt;210&gt; 6

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Primer oligonucleotide

&lt;400&gt; 6

tggctctcct caagcgatt c

21

&lt;210&gt; 7

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Primer oligonucleotide

&lt;400&gt; 7

actctctgca gtcaagcggc tca

23

&lt;210&gt; 8

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Primer oligonucleotide

&lt;400&gt; 8

cagtggatga accggctgat a

21

&lt;210&gt; 9

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Primer oligonucleotide

&lt;400&gt; 9

ggggcgacgg catcatgcaa acc

23

<210> 10  
<211> 23  
<212> DNA  
<213> Primer oligonucleotide

<400> 10  
caggacctgc agtctgtcaa agg 23

<210> 11  
<211> 17  
<212> DNA  
<213> Primer oligonucleotide

<400> 11  
cgggagagcc atagtgg 17

<210> 12  
<211> 19  
<212> DNA  
<213> Primer oligonucleotide

<400> 12  
agtaccacaa ggcctttcg 19

<210> 13  
<211> 21  
<212> DNA  
<213> Probe

<400> 13  
ctgcggaacc ggtgagtaca c 21

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
1 August 2002 (01.08.2002)

PCT

(10) International Publication Number  
**WO 2002/059321 A3**

(51) International Patent Classification<sup>7</sup>: C12N 15/40,  
15/51, 5/10, C07K 14/18, C12N 15/85, 7/00, G01N 33/50

(IT). PAONESSA, Giacomo [IT/IT]; IRBM, Via Pontina  
Km. 30, 600, I-00040 Pomezia (Rome) (IT).

(21) International Application Number:  
PCT/EP2002/000526

(74) Agent: THOMPSON, John; Merck & Co., Inc., European  
Patent Department, Terlings Park, Eastwick Road, Harlow,  
Essex CM20 2QR (GB).

(22) International Filing Date: 16 January 2002 (16.01.2002)

(81) Designated States (*national*): CA, JP, US.

(25) Filing Language: English

(84) Designated States (*regional*): European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE, TR).

(26) Publication Language: English

(30) Priority Data:  
60/263,479 23 January 2001 (23.01.2001) US

Published:

— with international search report  
— before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments

(71) Applicant (*for all designated States except US*): ISTI-  
TUTO DI RICERCHE DI BIOLOGIA MOLECO-  
LARE P. ANGELETTI SPA [IT/IT]; Via Pontina Km.  
30, 600, I-00040 Pomezia (Rome) (IT).

(88) Date of publication of the international search report:  
25 March 2004

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): DE FRANCESCO,  
Raffaele [IT/IT]; IRBM, Via Pontina Km. 30, 600, I-00040  
Pomezia (Rome) (IT). MIGLIACCIO, Giovanni [IT/IT];  
IRBM, Via Pontina Km. 30, 600, I-00040 Pomezia (Rome)

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

WO 2002/059321 A3

(54) Title: HEPATITIS C VIRUS REPLICONS AND REPLICON ENHANCED CELLS

(57) Abstract: The present invention features nucleic acid containing one or more adaptive mutations, and HCV replicon enhanced cells. Adaptive mutations are mutations that enhance HCV replicon activity. HCV replicon enhanced cells are cells having an increased ability to maintain an HCV replicon.

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/00526

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/40 C12N15/51 C12N5/10 C07K14/18 C12N15/85  
C12N7/00 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, Sequence Search, CHEM ABS Data, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|------------|---|-----------------------|
| X          | PATENT ABSTRACTS OF JAPAN<br>vol. 1995, no. 02,<br>31 March 1995 (1995-03-31)<br>& JP 06 319583 A (SOYAKU GIJUTSU<br>KENKYUSHO:KK),<br>22 November 1994 (1994-11-22)<br>abstract  | 1,9,19                |
| X          | -& DATABASE GENESEQ [Online]<br>E.B.I. Hinxton U.K.;<br>6 December 1995 (1995-12-06),<br>SOYAKU GIJUTSU KENKYUSHO KK.: "Hepatitis C<br>virus RNA helicase"<br>XP002245688<br>Database accession no. AAR68864<br>Ala at amino acid 1095<br>abstract<br><br>-----<br>-/-- | 1,9,19                |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

2 October 2003

Date of mailing of the international search report

12 JAN 2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Lorroy, 0

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/00526

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
|------------|--|-----------------------|
| X          | DATABASE EM VI [Online]<br>E.B.I. HINXTON U.K.;<br>27 April 2000 (2000-04-27),<br>NAGAYAMA K ET AL: "Hepatitis C virus<br>strain MD27 complete genome"<br>XP002245689<br>Database accession no. AF207768<br>abstract<br>-----  | 1,9,19                |
| A          | EP 1 043 399 A (BARTENSCHLAGER RALF DR)<br>11 October 2000 (2000-10-11)<br>cited in the application<br>claims 22-30<br>-----   |                       |
| A          | BLIGHT K J ET AL: "EFFICIENT INITIATION<br>OF HCV RNA REPLICATION IN CELL CULTURE"<br>SCIENCE, AMERICAN ASSOCIATION FOR THE<br>ADVANCEMENT OF SCIENCE,, US,<br>vol. 290, 8 December 2000 (2000-12-08),<br>pages 1972-1974, XP002951271<br>ISSN: 0036-8075<br>-----                                     |                       |
| A          | LOHMANN V ET AL: "REPLICATION OF<br>SUBGENOMIC HEPATITIS C VIRUS RNAS IN A<br>HEPATOMA CELLLINE"<br>SCIENCE, AMERICAN ASSOCIATION FOR THE<br>ADVANCEMENT OF SCIENCE,, US,<br>vol. 285, 2 July 1999 (1999-07-02), pages<br>110-113, XP000960693<br>ISSN: 0036-8075<br>cited in the application<br>----- |                       |



# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 02/00526

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---|---------------------|----------------------------|---------------------|
| JP 06319583                               | A | 22-11-1994          | NONE                       |                     |
| EP 1043399                                | A | 11-10-2000          | DE 19915178 A1             | 05-10-2000          |
|   |   |                     | AT 236988 T                | 15-04-2003          |
|   |   |                     | AU 2518000 A               | 19-10-2000          |
|   |   |                     | CA 2303526 A1              | 03-10-2000          |
|   |   |                     | DE 50001673 D1             | 15-05-2003          |
|   |   |                     | DK 1043399 T3              | 28-07-2003          |
|   |   |                     | EP 1043399 A2              | 11-10-2000          |
|   |   |                     | JP 2001017187 A            | 23-01-2001          |
|   |   |                     | US 6630343 B1              | 07-10-2003          |

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 02/00526

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1, 9, 19 all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1, 9, 19 (all partially)

A nucleic acid molecule comprising an altered HCV NS3 encoding region coding for one or more NS3 mutations, wherein at least one of said NS3 mutations, identified by reference to the amino acid sequence numbering of SeqIdNo.1, is amino acid 1095 being Ala; A nucleic acid molecule comprising an altered HCV NS3 encoding region coding for one or more NS3 mutations, wherein at least one of said NS3 mutations, identified by reference to the nucleotide sequence numbering of SeqIdNo.2, is nucleotide 3625 being cytosine; an expression vector comprising said nucleic acid molecule.

---

2. claims: 1, 9, 15-19 (all partially)

As for invention 1, wherein said at least one of said NS3 mutations is amino acid 1202 being Gly / nucleotide 3946 of SeqIdNo.2 being guanine / nucleotide 2330 of SeqIdNo.3 being guanine.

---

3. claims: 1, 9, 15-19 (all partially)

As for invention 1, wherein said at least one of said NS3 mutations is amino acid 1347 being Thr / nucleotide 4380 of SeqIdNo.2 being alanine / nucleotide 2764 of SeqIdNo.3 being alanine.

---

4. claims: 1-23 (all partially)

A nucleic acid molecule comprising an altered HCV NS5 encoding region coding for one or more NS5 mutations, wherein at least one of said NS5 mutations, identified by reference to the amino acid sequence numbering of SeqIdNo.1, is amino acid 2041 being Thr; A nucleic acid molecule comprising an altered HCV NS5 encoding region coding for one or more NS5 mutations, wherein at least one of said NS5 mutations, identified by reference to the nucleotide sequence numbering of SeqIdNo.2, is nucleotide 6463 being cytosine / nucleotide 4847 of SeqIdNo.3 being cytosine; said nucleic acid being an HCV replicon; expression vector and recombinant cell comprising said nucleic acid molecule.

---

5. claims: 1-23 (all partially)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

As for invention 4, wherein said at least one of said NS5 mutations is Lys insertion between residue 2039 and 2040 of SeqIdNo.1 / insertion of 3 adenine residues between nucleotide 6458 and 6459 of SeqIdNo.2 / insertion of 3 adenine residues after nucleotide 4843 of SeqIdNo.3

---

6. claims: 1-23 (all partially)

As for invention 4, wherein said at least one of said NS5 mutations is amino acid 2173 being Phe / nucleotide 6859 of SeqIdNo.2 being thymine or uracil / nucleotide 5243 of SeqIdNo.3 being thymine or uracil.

---

7. claims: 1-23 (all partially)

As for invention 4, wherein said at least one of said NS5 mutations is amino acid 2197 being Phe / nucleotide 6931 of SeqIdNo.2 being thymine or uracil / nucleotide 5315 of SeqIdNo.3 being thymine or uracil.

---

8. claims: 1-23 (all partially)

As for invention 4, wherein said at least one of said NS5 mutations is amino acid 2198 being Ser / nucleotide 6934 of SeqIdNo.2 being cytosine / nucleotide 5318 of SeqIdNo.3 being cytosine.

---

9. claims: 1-23 (all partially)

As for invention 4, wherein said at least one of said NS5 mutations is amino acid 2199 being Thr / nucleotide 6936 of SeqIdNo.2 being adenine / nucleotide 5320 of SeqIdNo.3 being adenine.

---

10. claims: 1-23 (all partially)

As for invention 4, wherein said at least one of said NS5 mutations is amino acid 2204 being Arg / nucleotide 6953 of SeqIdNo.2 being adenine or guanine / nucleotide 5337 of SeqIdNo.3 being adenine

---

11. claims: 1, 9, 15-19 (all partially)

A nucleic acid molecule comprising an altered EMCV IRES region wherein insertion of an extra adenosine at nucleotide 1736 of SeqIdNo.3 has occurred.

---

12. claims: 24-39 (all totally)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Method of making an HCV replicon enhanced cell comprising the steps of introducing and maintaining a HCV replicon in a cell, and curing said cell of said HCV replicon to produce said replicon enhanced cell; HCV replicon enhanced-cell obtained by said method; Method of measuring the ability of a compound to affect HCV activity using said cell.

---